(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 December 2004 (29.12.2004)

PCT

(10) International Publication Number
WO 2004/113522 A1

(51) International Patent Classification[7]: C12N 9/56, 9/50, 9/64, 9/00, A61K 38/48

(21) International Application Number: PCT/EP2004/051173

(22) International Filing Date: 18 June 2004 (18.06.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

03013819.2	18 June 2003 (18.06.2003)	EP
03025851.1	10 November 2003 (10.11.2003)	EP
03025871.9	11 November 2003 (11.11.2003)	EP
04003058.7	11 February 2004 (11.02.2004)	EP

(71) Applicant *(for all designated States except US)*: **DIREVO BIOTECH AG** [DE/DE]; Nattermannallee 1, 50667 Köln (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HAUPTS, Ulrich** [DE/DE]; c/o Direvo Biotech AG, 50667 Köln (DE). **KOLTERMANN, Andre** [DE/DE]; c/o Direvo Biotech AG, 50667 Köln (DE). **SCHEIDIG, Andreas** [DE/DE]; c/o Direvo Biotech AG, 50667 Köln (DE). **VOETSMEIER, Christian** [DE/DE]; c/o Direvo Biotech AG, 50667 Köln (DE). **KETTLING, Ulrich** [DE/DE]; c/o Direvo Biotech AG, 50667 Köln (DE).

(74) Agents: **HELBING, Jörg** et al.; Von Kreisler Selting Werner, Postfach 10 22 41, 50462 Köln (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NEW BIOLOGICAL ENTITIES AND THE PHARMACEUTICAL OR DIAGNOSTIC USE THEREOF

(57) Abstract: The present invention provides method for the treatment of a disease by applying a medicament comprising a protease with a defined specificity is capable to hydrolyze specific peptide bonds within a target substrate related to such disease. The proteases with such a defined specificity can further be used for related therapeutic or diagnostic purposes.

WO 2004/113522 A1

New Biological Entities and the Pharmaceutical or Diagnostic Use Thereof

The present invention provides methods for the treatment of a disease by applying a medicament comprising a protease with a defined specificity is capable to hydrolyze specific peptide bonds within a target substrate related to such disease. The proteases with such a defined specificity can further be used for related therapeutic or diagnostic purposes.

Background

Academic and industrial research continuously searches for functional proteins to be used as therapeutic, research, diagnostic, nutritional, personal care or industrial agents. Today, such functional proteins can be classified mainly into two categories: natural proteins and engineered proteins. Natural proteins, on the one hand, are discovered from nature, e.g. by screening natural isolates or by sequencing genomes from diverse species. Engineered proteins, on the other hand, are typically based on known proteins and are altered in order to acquire modified functionalities. The present invention discloses engineered proteins with novel functions as compared to the starting components. Such proteins are called NBEs (New Biologic Entities). The NBEs disclosed in the present invention are engineered enzymes with novel substrate specificities or fusion proteins of such engineered enzymes with other functional components.

Specificity is an essential element of enzyme function. A cell consists of thousands of different, highly reactive catalysts. Yet the cell is able to maintain a coordinated metabolism and a highly organized three-dimensional structure. This is due in part to the specificity of enzymes, i.e. the selective conversion of their respective substrates. Specificity is a qualitative and a quantitative property: the specificity of a particular enzyme can vary widely, ranging from just one particular type of target molecules to all molecular types with certain chemical substructures. In nature, the specificity of an organism's enzymes has been evolved to the particular needs of the organism. Arbitrary specificities with high value for therapeutic, research, diagnostic, nutritional or industrial applications are unlikely to be found in any organism's enzymatic repertoire due to the large

space of possible specificities. The only realistic way of obtaining such specificities is their generation de novo.

When comparing enzymes with binders, a paradigm of specificity is given by antibodies recognizing individual epitopes as small distinct structures within large molecules. The naturally occurring vast range of antibody specificities is attributed to the diversity generated by the immune system combined with natural selection. Several mechanisms contribute to the vast repertoire of antibody specificity and occur at different stages of immune response generation and antibody maturation (Janeway, C et al. (1999) Immunobiology, Elsevier Science Ltd., Garland Publishing, New York). Specifically, antibodies contain complementarity determining regions (CDRs) which interact with the antigen in a highly specific manner and allow discrimination even between very similar epitopes. The light as well as the heavy chain of the antibody each contribute three CDRs to the binding domain. Nature uses recombination of various gene segments combined with further mutagenesis in the generation of CDRs. As a result, the sequences of the six CDR loops are highly variable in composition and length and this forms the basis for the diversity of binding specificities in antibodies. A similar principle for the generation of a diversity of catalytic specificities is not known from nature.

Catalysis, i.e. the increase of the rate of a specific chemical reaction, is besides binding the most important protein function. Catalytic proteins, i.e. enzymes, are classified according to the chemical reaction they catalyze.

Transferases are enzymes transferring a group, for example, the methyl group or a glycosyl group, from one compound (generally regarded as donor) to another compound (generally regarded as acceptor). For example, glycosyltransferases (EC 2.4) transfer glycosyl residues from a donor to an acceptor molecule. Some of the glycosyltransferases also catalyze hydrolysis, which can be regarded as transfer of a glycosyl group from the donor to water. The subclass is further subdivided into hexosyltransferases (EC 2.4.1), pentosyltransferases (EC 2.4.2) and those transferring other glycosyl groups (EC 2.4.99, Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB)).

Oxidoreductases catalyze oxido-reductions. The substrate that is oxidized is regarded as hydrogen or electron donor. Oxidoreductases are classified as dehydrogenases, oxidases, mono- and dioxygenases. Dehydrogenases transfer hydrogen from a hydrogen donor to a hydrogen acceptor molecule. Oxidases react with molecular oxygen as hydrogen acceptor and produce oxidized products as well as either hydrogen peroxide or water. Monooxygenases transfer one oxygen atom from molecular oxygen to the substrate and one is reduced to water. In contrast, dioxygenases catalyze the insert of both oxygen atoms from molecular oxygen into the substrate.

Lyases calalyze elimination reactions and thereby generate double bonds or, in the reverse direction, catalyze the additions at double bonds. Isomerases catalyze intramolecular rearrangements. Ligases catalyze the formation of chemical bonds at the expense of ATP consumption.

Finally, hydrolases are enzymes that catalyze the hydrolysis of chemical bonds like C-O or C-N. The E.C. classification for these enzymes generally classifies them by the nature of the bond hydrolysed and by the nature of the substrate. Hydrolases such as lipases and proteases play an important role in nature as well in technical applications of biocatalysts. Proteases hydrolyse a peptide bond within the context of an oligo- or polypeptide. Depending on the catalytic mechanism proteases are grouped into aspartic, serin, cysteine, metallo- and threonine proteases (Handbook of proteolytic enzymes. (1998) Eds: Barret, A; Rawling, N.; Woessner, J.; Academic Press, London). This classification is based on the amino acid side chains that are responsible for catalysis and which are typically presented in the active site in very similar orientation to each other. The scissile bond of the substrate is brought into register with the catalytic residues due to specific interactions between the amino acid side chains of the substrate and complementary regions of the protease (Perona, J. & Craik, C (1995) *Protein Science*, 4, 337-360). The residues on the N- and C-terminal side of the scissile bond are usually called P_1, P_2, P_3 etc and P_1', P_2', P_3' and the binding pockets complementary to the substrate S_1, S_2, S_3 and S_1', S_2', S_3', respectively (nomenclature according to Schlechter & Berger, Biochem. Biophys. Res. Commun. 27 (1967) 157-162). The selectivity of proteases can vary widely from

being virtually nonselective – e.g. the Subtilisins – over a strict preference at the P_1 position – e.g. Trypsin selectively cutting on the C-terminal side of arginine or lysine residues – to highly specific proteases – e.g. human tissue-type plasminogen activator (t-PA) cleaving at the C-terminal side of the arginine in the sequence CPGRVVG (Ding, L et al. (1995) Proc. Natl. Acad. Sci. USA 92, 7627-7631; Coombs, G et al. (1996) *J. Biol. Chem.* 271, 4461-4467).

The specificity of proteases, i.e. their ability to recognize and hydrolyze preferentially certain peptide substrates, can be expressed qualitatively and quantitatively. Qualitative specificity refers to the kind of amino acid residues that are accepted by a protease at certain positions of the peptide substrate. For example, trypsin and t-PA are related with respect to their qualitative specificity, since both of them require at the P_1 position an arginine or a similar residue. On the other hand, quantitative specificity refers to the relative number of peptide substrates that are accepted as substrates by the protease, or more precisely, to the relative k_{cat}/k_M ratios of the protease for the different peptides that are accepted by the protease. Proteases that accept only a small portion of all possible peptides have a high specificity, whereas the specificity of proteases that, as an extreme, cleave any peptide substrate would theoretically be zero.

Comparison of the primary, secondary as well as the tertiary structure of proteases (Fersht, A., Enzyme Structure and Mechanism, W. H. Freeman and Company, New York, 1995) allows identification of classes showing a high degree of conservation (Rawlings, N.D. & Barrett, A.J. (1997) In: *Proteolysis in Cell Functions* Eds. Hopsu-Havu,V.K.; Järvinen,M.; Kirschke,H, pp. 13-21, IOS Press, Amsterdam). A widely accepted scheme for protease classification has been proposed by Rawlings & Barrett (Handbook of proteolytic enzymes. (1998) Eds: Barret, A; Rawling, N.; Woessner, J.; Academic Press, London). For example, the serine proteases family can be subdivided into structural classes with chymotrypsin (class S1), subtilisin (class S8) and carboxypeptidase (class SC) folds, each of which includes nonspecific as well as specific proteases (Rawlings, N.D. & Barrett, A.J. (1994) *Methods Enzymol.* 244, 19-61). This applies to other protease families analogously. An additional distinction can be made according to the relative location of the cleaved bond in the substrate. Carboxy- and

aminopeptidases cleave amino acids from the C- and N-terminus, respectively, while endopeptidases cut anywhere along the oligopeptide.

Many applications would be conceivable if enzymes with a basically unlimited spectrum of specificities were available. However, the use of such enzymes with high, low or any defined specificity is currently limited to those which can be isolated from natural sources. The field of application for these enzymes varies from therapeutic, research, diagnostic, nutritional to personal care and industrial purposes.

Enzyme additives in detergents have come to constitute nearly a third of the whole industrial enzyme market. Detergent enzymes include proteinases for removing organic stains, lipases for removing greasy stains, amylases for removing residues of starchy foods and cellulases for restoring of smooth surface of the fiber. The best-known detergent enzyme is probably the nonspecific proteinase subtilisin, isolated from various *Bacillus* species.

Starch enzymes, such as amylases, occupy the majority of those used in food processing. While starch enzymes include products that are important for textile desizing, alcohol fermentation, paper and pulp processing, and laundry detergent additives, the largest application is for the production of high fructose corn syrup. The production of corn syrup from starch by means of industrial enzymes was a successful alternative to acid hydrolysis.

Apart from starch processing, enzymes are used for an increasing range of applications in food. Enzymes in food can improve texture, appearance and nutritional value or may generate desirable flavours and aromas. Currently used food enzymes in bakery are amylase, amyloglycosidases, pentosanases for breakdown of pentosan and reduced gluten production or glucose oxidases to increase the stability of dough. Common enzymes for dairy are rennet (protease) as coagulant in cheese production, lactase for hydrolysis of lactose, protease for hydrolysis of whey proteins or catalase for the removel of hydrogen peroxides. Enzymes used in brewing process are the above named amylases, but also cellulases or proteases to clarify the beer from suspended proteins. In wines and fruit juices, cloudiness is more commenly caused by starch and pectins so that

amylases and pectinases increase yield and clarification. Papain and other proteinases are used for meat tenderizing.

Enzymes have also been developed to aid animals in the digestion of feed. In the western hemisphere, corn is a major source of food for cattle, swine, and poultry. In order to improve the bioavailability of phosphate from corn, phytase is commonly added (Wyss, M. et al., Biochemical characterization of fungal phytases (myo-inositol hexakisphosphate phosphohydrolases); Catalytic properties. Applied & Environmental Microbiology 65, 367-373 (1999)). Moreover, phytate hydrolysis has been shown to bring about improvements in digestibility of protein and absorption of minerals such as calcium (Bedford, M.R. & Schulze, H., Exogenous Enzymes for Pigs and Poultry [Review]. Nutrition Research Reviews 11, 91-114 (1998)). Another major feed enzyme is xylanase. This enzyme is particularly useful as a supplement for feeding stuff comprising more than about 10% of wheat barley or rye, because of their relatively high soluble fiber content. Xylanases cause two important actions: reduction of viscosity of the intestinal contents by hydrolyzing the gel-like high molecular weight arabinoxylans in feed (Murphy, T et al., Effect of range of new xylanases on in vitro viscosity and on performance of broiler diets. British Pultry Science 44, S16-S18 (2003)) and break down of polymers in cell wallswhich improve the bioavailability of protein and starch.

Biotech research and development laboratories routinely use special enzymes in small quantities along with many other reagents. These enzymes create a significant market for various enzymes. Enzymes like alkaline phosphatase, horseradish peroxidase and luciferase are only some examples. Thermostable DNA polymerases like Taq polymerase or restriction endonucleases revolutionized laboratory work.

The use of enzymes in the diagnosis of disease is another important benefit derived from the intensive research in biochemistry. Within the recent past few years that interest in diagnostic enzymology has increased and there are still large areas of medical research in which the diagnostic potential of enzyme reactions has not been explored at all. Common enzymes used for clinical diagnosis are acid phosphatase, alanine aminotransferase, alkaline phosphatase,

amylase, angiotensin converting enzymes, aspartate aminotransferase, cholinesterase, creatinine kinase, gamma glutamyltransferase, lactate dehydrogenaseor rennin.

Therapeutic enzymes are a particular class of drugs, categorized by the FDA as biologicals, with a lot of advantages compared to other, especially non-biological pharmaceuticals. Examples for successful therapeutic enzymes are human clotting factors like factor VIII and factor IX for human treatment. In addition, digestive enzymes are used for various deficiencies in human digestive processes. Other examples are t-PA and streptokinase for the treatment of cardiovascular disease, beta-glucocerebrosidase for the treatment of Type I Gaucher disease, L-asparaginase for the the treatment of acute lymphoblastic leukemia and DNAse for the treatment of cystic fibrosis. An important issue in the application of proteins as therapeutics is their potential immunogenicity. To reduce this risk, one would prefer enzymes of human origin, which narrows down the set of available enzymes. The provision of designed enzymes, preferably of human origin, with novel, tailor-made specificities would allow the specific modification of target substrates at will, while minimizing the risk of immunogenicity. A further advantage of highly specific enzymes as therapeutics would be their lower risk of side effects. Due to the limited possibility of specific interactions between a small molecule and a protein, binding to non-target proteins and therefore side effects are quite common and often cause termination of an otherwise promising lead compound. Specific enzymes, on the other hand, provide many more contact sites and mechanisms for substrate discrimination and therefore enable a higher specificity and thereby less side activities.

Proteases represent an important class of therapeutic agents (*Drugs of today*, 33, 641-648 (1997)). However, currently the therapeutic protease is usually a substitute for insufficient acitivity of the body's own proteases. For example, factor VII can be administered in certain cases of coagulation deficiencies of bleeders or during surgery (Heuer L.; Blumenberg D. (2002) *Anaesthesist* 51:388). Tissue-type plasminogen activator (t-PA) is applied in acute cardiac infarction, initializing the dissolution of fibrin clots through specific cleavage and activation of plasminogen (Verstraete, M. et al. (1995) *Drugs*, 50, 29-41). So far

a protease with taylor-made specificity is generated to provide a therapeutic agent that specifically activates or inactivates a disease related target protein.

Monoclonal antibodies represent another important biological class of substances with therapeutic capabilities. One of the main antibody targets are tumor necrosis factors (TNFs) which belong to the family of cytokines. TNFs play a major role in the inflammation process. As homotrimers they could bind to receptors of nearly every cell. They activate a multiplicity of cellular genes, multiple signal transduction mechanisms, kinases and transcription factors. The most important TNFs are TNF-alpha and TNF-beta. TNF-alpha is produced by macrophages, monocytes and other cells. TNF-alpha is an inflammation mediator. Therefore, research of the last decade has been focused on TNF-alpha inhibitors like monoclonal antibodies as possible therapeutics for different therapeutic indications like Rheumatoid Arthritis, Crohn´s disease or Psoriasis (Hamilton et al. (2000) Expert Opin Pharmacother, 1 (5): 1041-1052). One of the major disadvantages of monoclonal antibodies are their high costs, so that new biological alternatives are of great importance.

There are a lot of examples for engineered enzymes in literature. Fulani et al. (Fulani F. et al. (2003) *Protein Engineering* 16, 515-519) describe a rhodanase (thiosulfat:cyanide sulfurtransferase) from Azotobacter vinelandii which has a catalytic domain structurally related to catalytic subunit of Cdc25 phosphatase enzymes. The difference in catalytic mechanism depends on the different size of the active site. Both rhodanase and phosphatase are highly specific on different substrates (sulfate vs. phosphate). The catalytic mechanism of the rhodanase could be shifted towards serine/threonine phosphatase by single-residue insertion. Therefore, Fulani et al. give a single example for the change of a catalytic mechanism by structural comparison and sequence alignment of naturally known enzymes from different enzyme classes but lack an indication of how to generate a user-definable substrate specificity while keeping the same catalytic mechanism.

The thioredoxin reductase described by Briggs et al. (WO 02/090300 A2) has an altered cofactor specificity which preferably binds NADPH compared to NADH. Thus, both enzymes, the starting point as well as the resulting engineered

enzyme are highly specific towards different substrates. The methods to achieve such an altered substrate specificity are either computational processing methods or sequence alignments of related proteins to define variable and conserved residues. They all have in common that they are based on the comparison of structures and sequences of proteins with known specificities followed by the transfer of the same to another backbone.

There are other examples of specificity-engineered enzymes and, in particular, of proteases which have been published in the literature. None of these examples, however, provides a means for generating novel specificites compared to the specificity of the starting material used within the described methods. The methods range from structure-directed single point mutations (Kurth, T. et al. (1998) *Biochemistry* 37, 11434-11440; Ballinger, M et al. (1996) *Biochemistry*, 35:13579-13585), exchange of surface loops between two specific proteases (Horrevoets et al. (1993) *J. Biol. Chem*. 268, 779-782), to random mutagenesis either regio-selectively or across the whole gene combined with in-vitro or in-vivo selection (Sices, H. & Kristie, T. (1998) *Proc. Natl. Acad. Sci. USA*, 95, 2828-2833).

The rational design of protease specificity is limited to very few examples. This approach is severely limited by the insufficient understanding of the complexities that govern folding and dynamics as well as structure-function relationships in proteins (Corey, M.J. & Corey, E. (1996) *Proc. Natl. Acad. Sci. USA*, 93:11428-11434). It is therefore difficult to alter the primary amino acid sequence of a protease in order to change its activity or specificity in a predictive way. In a successful example, Kurth et al. engineered trypsin to show a preference for a dibasic motive (Kurth, T. et al. (1998) *Biochemistry*, 37:11434-11440). In another example, Hedstrom et al. converted the S_1 substrate specificity of trypsin to that of chymotrypsin (Hedstrom, L. et al. (1992) Science, 255:1249-1253). This is an example where a known property was transferred from one backbone to another.

Ballinger et al. (WO 96/27671) describe subtilisin variants with combination mutations (N62D/G166D, and optionally Y104D) having a shift of substrate specificity towards peptide or polypeptide substrates with basic amino acids at

the P1, P2 and P4 positions of the substrate. Suitable substrates of the variant subtilisin were revealed by sorting a library of phage particles (substrate phage) containing five contiguous randomized residues. These subtilisin variants are useful for cleaving fusion proteins with basic substrate linkers and processing hormones or other proteins (in vitro or in vivo) that contain basic cleavage sites. The problems associated with rational redesign of enzymes can partially be overcome by directed evolution (as disclosed in PCT/EP03/04864). These studies can be classified by their expression and selection systems. Genetic selection means to produce inside an organism an enzyme, e.g. a protease, which is able to cleave a precursor protein which in turn results in an alteration of the growth behavior of the producing organism. From a population of organisms with different proteases those can be selected which have an altered growth behavior. This principle was for example reported by Davis et al. (US 5258289, WO 96/21009). The production of a phage system is dependent on the cleavage of a phage protein which only can be activated in the presence of a proteolytic enzyme which is able to cleave the phage protein. Other approaches use a reporter system which allows a selection by screening instead of a genetic selection, but also cannot overcome the intrinsic insufficiency of the intracellular characterization of enzymes.

Systems to generate enzymes with altered sequence specificities with self-secreting enzymes are also reported. Duff et al. (WO 98/11237) describe an expression system for a self-secreting protease. An essential element of the experimental design is that the catalytic reaction acts on the protease itself by an autoproteolytic processing of the membrane-bound precursor molecule to release the matured protease from the cellular membrane into the extracellular environment. Therefore, a fusion protein must be constructed where the target peptide sequence replaces the natural cleavage site for autoproteolysis. Limitations of such a system are that positively identified proteases will have the ability to cleave a certain amino acid sequence but they also may cleave many other peptide sequences. Therefore, high substrate specificity cannot be achieved. Additionally, such a system is not able to control that selected proteases cleave at a specific position in a defined amino acid sequence and it does not allow a precise characterization of the kinetic constants of the selected proteases (k_{cat}, K_M).

A method has been described that aims at the generation of new catalytic activities and specificities within the α/ß-barrel proteins (WO 01/42432; Fersht et al, Methods of producing novel enzymes; Altamirano et al. (2000) *Nature* 403, 617-622). The α/ß-barrel proteins comprise a large superfamily of proteins accounting for a large fraction of all known enzymes. The structure of the proteins is made from a/ß-barrel surrounded by α-helices. The loops connecting ß-strands and helices comprise the so-called lid-structure including the acitve site residues. The method is based on the classification of α/ß-barrel proteins into two classes based on the catalytic lid structure. An extensive comparison of α/ß-barrel protein structures led the authors to the conclusion that the substrate binding and specificity is primarily defined by the barrel structure while the specificity of the chemical reaction resides within the loops. It is suggested that barrels and lid structures from different enzymes can be combined to generate new enzymatic activities and to provide a starting point to fine tune the properties by targeted or randomized mutagenesis and selection. The method does not provide for the generation of user-defined specificity.

In summary, it is clear that there are many possible applications in the fields of therapeutics, research and diagnostics, industrial enzymes, food and feed processing, cosmetics and other areas that would become possible by the availability of enzymes with a novel substrate specificity. However, only a limited number of specific enzymes has been identified from natural sources so far. Methods of rational design to modify, alter, convert or transfer sequence specificity as well as random approaches described above did not enable the generation of a novel and user-definablespecificity that was not present in the employed starting material.

Therefore, none of the currently available methods can provide enzymes with a novel and user-defined sequence specificity. In contrast, the current invention provides such enzymes as well as methods for generating them.

Summary of the Invention

The objective of the present invention is to provide a method for the treatment of a disease by applying a medicament comprising a protease. Further the

present invention provides engineered proteins with novel functions that do not exist in the components used for the engineering of such proteins. In particular, the invention provides enzymes with user-definable specificities. User-definable specificity means that enzymes are provided with specificities that do not exist in the components used for the engineering of such enzymes. The specificities can be chosen by the user so that one or more intended target substrates are preferentially recognised and converted by the enzymes. Furthermore, the invention provides enzymes that possess essentially identical sequences to human proteins but have different specificities. In a particular embodiment, the invention provides proteases with user-definable specificities.

Furthermore, the present invention is directed to engineered enzymes which are fused to one or more further functional components. These further components can be proteinacious components which preferably have binding properties and are of the group consisting of substrate binding domains, antibodies, receptors or fragments thereof. Furthermore, these further components can be further functional components, preferably being selected from the group consisting of polyethylenglycols, carbohydrates, lipids, fatty acids, nucleic acids, metals, metal chelates, and fragments or derivatives thereof. The resulting fusion proteins are understood as enzymes with user-definable specificities within the present invention.

Besides, the invention is directed to the application of such enzymes with novel, user-definable specificities for therapeutic, research, diagnostic, nutritional, personal care or industrial purposes. Moreover, the invention is directed to a method for generating engineered enzymes with user-definable specificities. In particular, the invention is directed to generate enzymes that possess essentially identical sequences to human enzymes but have different specificities.

This problem has been solved by the embodiments of the invention specified in the description below and in the claims. The present invention is thus directed to (1) the use of a protease with defined specificity for a target substrate for preparing a medicament for the treatment of a specific disease related to said target substrate,

(2) an engineered enzyme with defined specificity characterized by the combination of the following components,:
(a) a protein scaffold which catalyzes at least one chemical reaction on at least one substrate, and
(b) one or more specificity determining regions (SDRs) located at sites in the protein scaffold that enable the resulting engineered protein to discriminate between at least one target substrate and one or more different substrates, and wherein the SDRs are essentially synthetic peptide sequences;
(3) the use of an engineered enzyme as defined in (2) above for therapeutic, research, diagnostic, nutritional, personal care or industrial purposes, preferably for the use as defined in (1) above;
(4) a method for generating engineered enzymes as defined in (2) above having specificities towards target substrates, such specificities not being present in the individual starting components, comprising at least the following steps:
(a) providing a protein scaffold which catalyzes at least one chemical reaction on at least one substrate,
(b) generating a library of engineered enzymes by combining the protein scaffold from step (a) with fully or partially random peptide sequences at sites in the protein scaffold that enable the resulting engineered enzyme to discriminate between at least one target substrate and one or more different substrates, and
(c) selecting out of the library of engineered enzymes generated in step (b) one or more enzymes that have specificities towards at least one target substrate;
(5) a fusion protein which is comprised of at least one engineered enzyme as defined in (2) above and at least one further component, preferably the at least one further component having binding properties and more preferably being selected from the group consisting of antiboides, binding domains, receptors, and fragments thereof;
(6) a composition or pharmaceutical composition comprising one or more engineered enzymes as defined in (2) above or a fusion protein as defined in (5) above, said pharmaceutical composition may optionally comprise an acceptable carrier, excipient and/or auxiliary agent;
(7) a DNA encoding the engineered enzyme as defined in (2) above;
(8) a vector comprising the DNA as defined in (7) above;

(9) a host cell or transgenic organism being transformed/transfected with a vector as defined in (8) above and/or containing the DNA as defined in (7) above; and
(10) a method for producing the engineered enzyme of (2) above comprising culturing a cell or organism as defined in (8) above and isolating the enzyme from the culture broth.

Brief description of the Figures
The following figures are provided in order to explain further the present invention in supplement to the detailed description:

Figure 1 illustrates the three-dimensional structure of human trypsin I with the active site residues shown in "ball-and-stick" representation and with the marked regions indicating potential SDR insertion sites.

Figure 2 shows the alignment of the primary amino acid sequence of three members of the serine protease class S1 family: human trypsin I, human alpha-thrombin and human enteropeptidase (see also SEQ ID NOs: 1, 5 and 6).

Figure 3 illustrates the three-dimensional structure of subtilisin with the active site residues being shown in "ball-and-stick" representation and with the numbered regions indicating potential SDR insertion sites.

Figure 4 shows the alignment of the primary amino acid sequences of four members of the serine protease class S8 family: subtilisin E, furin, PC1 and PC5 (see also SEQ ID NOs: 7-10).

Figure 5 illustrates the three-dimensional structure of pepsin with the active site residues being shown in "ball-and-stick" representation and with the numbered regions indicating potential SDR insertion sites.

Figure 6 shows the alignment of the primary amino acid sequences of three members of the A1 aspartic acid protease family: pepsin, ß-secretase and cathepsin D (see also SEQ ID NOs: 11-13).

Figure 7: illustrates the three-dimensional structure of caspase 7 with the active site residues being shown in "ball-and-stick" representation and with the numbered regions indicating potential SDR insertion sites.

Figure 8: shows the primary amino acid sequence of caspase 7 as a member of the cysteine protease class C14 family (see also SEQ ID NO: 14).

Figure 9 depicts schematically the third aspect of the invention.

Figure 10 shows a Western blot analysis of a culture supernatant of cells expressing variants of human trypsin I with SDR1 and SDR2, compared to negative controls.

Figure 11 shows the time course of the proteolytic cleavage of a target substrate by human trypsin I.

Figure 12 shows the relative activities of three variants of inventive engineered proteolytic enzymes in comparison with human trypsin I on two different peptide substrates.

Figure 13 shows the relative specificities of human trypsin I and variants of inventive engineered proteolytic enzymes with one or two SDRs, respectively.

Figure 14: shows the relative specificities of human trypsin I and of variants of inventive engineered proteolytic enzymes being specific for human TNF-alpha with this scaffold on peptides with a target sequence of human TNF-alpha.

Figure 15: shows the reduction of cytotoxicity induced by TNF-alpha when incubating the TNF-alpha with concentrated supernatant from cultures expressing the inventive engineered proteolytic enzymes being specific for human TNF-alpha.

Figure 16: shows the reduction of cytotoxicity induced by TNF-alpha when incubating the TNF-alpha with purified inventive engineered proteolytic enzyme being specific for human TNF-alpha.

Figure 17: compares the activity of inventive engineered proteolytic enzymes being specific for human TNF-alpha with the activity of human trypsin I on two protein substrates: (a) human TNF-alpha; (b) mixture of human serum proteins.

Figure 18: showes the specific activity of an inventive engineered proteolytic enzyme with specificity for human VEGF.

Definitions

In the framework of the present invention the following terms and definitions are used.

The term "protease" means any protein molecule that is capable of hydrolysing peptide bonds. This includes naturally-occurring or artificial proteolytic enzymes, as well as variants thereof obtained by site-directed or random mutagenesis or any other protein engineering method, any active fragment of a proteolytic enzyme, or any molecular complex or fusion protein comprising one of the aforementioned proteins. A "chimera of proteases" means a fusion protein of two or more fragments derived from different parent proteases.

The term "substrate" means any molecule that can be converted catalytically by an enzyme. The term "peptide substrate" means any peptide, oligopeptide, or protein molecule of any amino acid composition, sequence or length, that contains a peptide bond that can be hydrolyzed catalytically by a protease. The peptide bond that is hydrolyzed is referred to as the "cleavage site". Numbering of positions in the substrate is done according to the system introduced by Schlechter & Berger (Biochem. Biophys. Res. Commun. 27 (1967) 157-162). Amino acid residues adjacent N-terminal to the cleavage site are numbered P_1, P_2, P_3, etc., whereas residues adjacent C-terminal to the cleavage site are numbered P_1', P_2', P_3', etc.

The term "target substrate" describes a user-defined substrate which is specifically recognized and converted by an enzyme according to the invention. The term "target peptide substrate" describes a user-defined peptide substrate.

The term "target specificity" describes the qualitative and quantitative specificity of an enzyme that is capable of recognizing and converting a target substrate. Catalytic properties of enzymes are expressed using the kinetic parameters "K_M" or "Michaelis Menten constant", "k_{cat}" or "catalytic rate constant", and "k_{cat}/K_M" or "catalytic efficiency", according to the definitions of Michaelis and Menten (Fersht, A., Enzyme Structure and Mechanism, W. H. Freeman and Company, New York, 1995). The term "catalytic activity" describes quantitatively the conversion of a given substrate under defined reaction conditions.

The term "specificity" means the ability of an enzyme to recognize and convert preferentially certain substrates. Specificity can be expressed qualitatively and quantitatively. "Qualitative specificity" refers to the chemical nature of the substrate residues that are recognized by an enzyme. "Quantitative specificity" refers to the number of substrates that are accepted as substrates. Quantitative specificity can be expressed by the term s, which is defined as the negative logarithm of the number of all accepted substrates divided by the number of all possible substrates. Proteases, for example, that accept preferantially a small portion of all possible peptide substrates have a "high specificity". Proteases that accept almost any peptide substrate have a "low specificity". Definitions are made in accordance to WO 03/095670 which is therefore incorporated by reference.Proteases with very low specificity are also referred to as "unspecific proteases". The term "defined specificity" refers to a certain type of specificity, i.e. to a certain target subtrate or a set of certain target substrates that are preferentially converted versus other substrates.

The term "engineered" in combination with the term "enzyme" describes an enzyme that is comprised of different components and that has features not being conferred by the individual components alone.

The term "protein scaffold" or "scaffold protein" refers to a variety of primary, secondary and tertiary polypeptide structures.

The term "peptide sequence" indicates any peptide sequence used for insertion or substitution into or combination with a protein scaffold. Peptide sequences are usually obtained by expression from DNA sequences which can be synthesized

according to well-established techniques or can be obtained from natural sources. Insertion, substitution or combination of peptide sequences with the protein scaffold are generated by insertion, substitution or combination of oligonucleotides into or with a polynucleotide encoding the protein scaffold. The term "synthetic" in combination with the term "peptide sequence" refers to peptide sequences that are not present in the protein scaffold in which the peptide sequences are inserted or substituted or with which they are combined.

The term "components" in combination with the term "engineered enzyme" refers to peptide or polypeptide sequences that are combined in the engineering of such enzymes. Such components may among others comprise one or more protein scaffolds and one or more synthetic peptide sequences. The term "library of engineered enzymes" describes a mixture of engineered enzymes, whereby every single engineered enzyme is encoded by a different polynucleotide sequence. The term "gene library" indicates a library of polynucleotides that encodes the library of engineered enzymes. The term "SDR" or "Specificity determining region" refers to a synthetic peptide sequence that provides the defined specificity when combined with the protein scaffold at sites that enable the resulting enzymes to discriminate between the target substrate and one or more other substrates. Such sites are termed "SDR sites".

The terms "tertiary structure similar to the structure of" and "similar tertiary structure" in combination with the terms "enzyme" or "protein" refer to proteins in which the type, sequence, connectivity and relative orientation of the typical secondary structural elements of a protein, e.g. alpha-helices, beta-sheets, beta-turns and loops, are similar and the proteins are therefore grouped into the same structural or topological class or fold. This includes proteins that have altered, additional or deleted structural elements of any type but otherwise unchanged topology. Examples of such structural classes are the TNF superfamily, the S1 fold or the S8 fold within the serine proteases, the GPCRs, or the α/β-barrel fold.

The term "positions that correspond structurally" indicates amino acids in proteins of similar tertiary structure that correspond structurally to each other, i.e. they are usually located within the same structural or topological element of the structure. Within the structural element they possess the same relative

positions with respect to beginning and end of the structural element. If, e.g. the topological comparison of two proteins reveals two structurally corresponding sequences of different length, then amino acids within, e.g. 20% and 40% of the respective region lengths, correspond to each other structurally.

The term "library of engineered enzymes" of the present invention refers to a multiplicity of enzymes or enzyme variants, which may exist as a mixture or in isolated form.

Amino acids residues are abbreviated according to the following Table 1 either in one- or in three-letter code.

Table 1: Amino acid abbreviations

Abbreviations		Amino acid
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophane
Y	Tyr	Tyrosine

Detailed description of the invention

The present invention provides engineered proteins with novel functions. In particular, the invention provides enzymes with user-definable specificities. In a particular embodiment, the invention provides proteases with user-definable specificities. Besides, the invention provides applications of such enzymes with novel, user-definable specificities for therapeutic, research, diagnostic, nutritional, personal care or industrial purposes. Moreover, the invention provides a method for generating enzymes with specificities that are not present in the components used for the engineering of such enzymes. In particular, the invention is directed to the generation of enzymes that have sequences that are essentially identical to mammalian, especially human enzymes but have different specificities . Moreover, the invention provides libraries of specific engineered enzymes with corresponding specificities encoded genetically, a method for the generation of libraries of specific engineered enzymes with corresponding specificities encoded genetically, and the application of such libraries for technical, diagnostic, nutritional, personal care or research purposes.

A first aspect of the invention is directed to the application of engineered enzymes with specificities for therapeutic, research, diagnostic, nutritional, personal care or industrial purposes. The application comprises at least the following steps:

(a) identification of a target peptide substrate whose hydrolysis has a positive effect in connection with the intended purpose, such as curing a disease, diagnosing a disease, processing of ingredients for human or animal nutrition, or other technical processes;

(b) provision of an engineered enzyme, the enzyme being specific for the target peptide identified in step (a); and

(c) use of the enzyme as provided in step (b) for the intended purpose.

In a first variant of this aspect of the invention, the engineered enzyme is used as a therapeutic means to inactivate a disease-related target substrate. This application comprises at least the following steps:

(a) identification of a target substrate whose function is connected to a disease and whose inactivation has a positive effect in connection with the disease, and determination of a target site within the target substrate

characterized by the fact that modification at the target site leads to the inactivation of the target substrate;

(b) provision of an engineered enzyme, the enzyme being specific for the target site identified in step (a); and

(c) use of the enzyme for the inactivation of the target substrate inside or outside the human body.

Preferably, the scaffold is a protease and the modification is hydrolysis of a target site in a protein target. Preferably, the hydrolysis leads to the activation or inactivation of the peptide or protein target. Potential peptide or protein targets include soluble proteins, in particular cytokines, such as proteins of the TNF-superfamily, interleukines, interferons, chemokines and growth factors; hormones; toxins; enzymes, such as oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases; structural proteins, such as collagen; and immunoglobulins; or membrane associated proteins, in particular single pass transmembrane proteins; multipass transmembrane proteins, such as G-protein coupled receptors, ion channels and transporters; lipid-anchored membrane proteins and GPI-anchored membrane proteins.

In a first embodiment of this variant the engineered enzyme is a protease and is capable of hydrolysing human tumor necrosis factor-alpha (hTNF-α). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, rheumatoid arthritis, inflammatory bowel diseases, psoriasis, Crohn's disease, Ulcerative colitis, diabetes type II, classical Hodgkin's Lymphoma (cHL), Grave's disease, Hashimoto's thyroiditis, Sjogren's Syndrome, systemic lupus erythematosus, multiple sclerosis, Systemic inflammatory response syndrome (SIRS) which leads to distant organ damage and multiple organ dysfunction syndrome (MODS), eosinophilia, neurodegenerative disease, stroke, closed head injury, encephalitis, CNS disorders, asthma, rheumatoid arthritis, sepsis, vasodilation, intravascular coagulation and multiple organ failure, as well as other diseases connected with hTNF-α. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hTNF-α (SEQ ID NO:96). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 31/32, 32/33, 44/45, 45/46, 87/88, 128/129, 130/131, 140/141 and/or 141/142 (most

preferred between positions 31/32, 32/33 and/or 45/46) in hTNF-α, or a peptide bond in proximity to these positions in hTNF-α, or peptide bonds in protein targets related to hTNF-α between positions having structural homology or sequence homology to these positions. In this embodiment it is most preferred that the protease has the a sequence shown in SEQ ID NO:74, SEQ ID NO:75 and is capable of hydrolysing hTNF-α at positions 31/32 and/or 32/33.

In a second embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Tumor necrosis factor ligand superfamily member 5 (hCD40-L). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, systemic lupus erythematosus and classical Hodgkin's Lymphoma (cHL), as well as other diseases connected with hCD40-L. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hCD40-L (SEQ ID NO:143). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 133/134, 145/146, 165/166, 200/201, 201/202, 207/208 and/or 216/217 (most preferred between positions 133/134, 165/166, 201/202 and/or 216/217) in hCD40-L, or a peptide bond in proximity to these positions in hCD40-L, or peptide bonds in protein targets related to hCD40-L at positions having structural homology or sequence homology to these positions.

In a third embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Macrophage migration inhibitory factor (hMIF). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, inflammatory diseases, as well as other diseases connected with hMIF. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hMIF (SEQ ID NO:109). More preferably said engineered or said fusion protein is capable of hydrolysing the peptide bonds between positions 16/17, 44/45, 66/67, 73/74, 77/78, 88/89, 92/93 and/or 100/101 (most preferred between positions 16/17 and/or 92/93) in hMIF, or a peptide bond in proximity to these positions in hMIF, or peptide bonds in protein targets related to hMIF at positions having structural homology or sequence homology to these positions.

In a fourth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin-1 beta precursor (hIL-1 beta). The enzymes or

the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, diabetes, brain inflammation in cancer, arthritis, autoimmune and inflammatory diseases, as well as other diseases connected with hIL-1 beta. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-1 beta (SEQ ID NO:112). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 24/25, 35/36, 46/47, 54/55, 74/75, 75/76, 76/77, 77/78, 86/87, 88/89, 93/94, 94/95, 97/98 and/or 150/151 (most preferred between positions 35/36, 75/76, 76/77, 88/89, 93/94, 94/95 and/or 150/151) in hIL-1 beta, or a peptide bond in proximity to these positions in hIL-1 beta, or peptide bonds in protein targets related to hIL-1 beta at positions having structural homology or sequence homology to these positions.

In a fifth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin 2 (hIL-2). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, T-cell leukemia and hairy cell leukemia, Crohn's disease, Ulcerative colitis, Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, systemic lupus erythematosus, multiple sclerosis, asthma and chronic obstructive pulmonary and classical Hodgkin's Lymphoma (cHL), as well as other diseases connected with hIL-2. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-2 (SEQ ID NO:99). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 20/21, 32/33, 38/39, 43/44, 45/46 48/49, 49/50, 54/55, 64/65, 76/77, 83/84, 84/85, 107/108, 109/110 and/or 120/121 (most preferred between positions 109/110) in hIL-2, or a peptide bond in proximity to these positions in hIL-2, or peptide bonds in protein targets related to hIL-2 at positions having structural homology or sequence homology to these positions.

In a sixth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin 3 (hIL-3). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, classical Hodgkin's Lymphoma (cHL) and eosinophilia, as well as other diseases connected withh IL-3. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-3 (SEQ ID NO:148). More preferably said enzyme or said fusion protein is capable of hydrolysing the

peptide bonds between positions 21/22, 28/29, 36/37, 44/45, 46/47, 51/52, 63/64, 66/67, 79/80, 94/95, 101/102, 108/109 and/or 109/110 (most preferred between positions 21/22, 28/29, 46/47, 63/64, 66/67, 79/80 and/or 101/102) in hIL-3, or a peptide bond in proximity to these positions in hIL-3, or peptide bonds in protein targets related to hIL-3 at positions having structural homology or sequence homology to these positions.

In a seventh embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin 4 (hIL-4). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, classical Hodgkin's Lymphoma (cHL), Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, Asthma, chronic obstructive pulmonary disease and allergic inflammatory reactions, as well as other diseases connected with hIL-4. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-4 (SEQ ID NO:118). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 4/5, 12/13, 31/32, 37/38, 61/62, 62/63, 64/65, 91/92, 102/103, 121/122 and/or 126/127 (most preferred between positions 4/5, 61/62, 62/63, 64/65 and/or 121/122) in hIL-4, or a peptide bond in proximity to these positions in hIL-4, or peptide bonds in protein targets related to hIL-4 at positions having structural homology or sequence homology to these positions.

In a eighth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin-5 (hIL-5). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, classical Hodgkin's Lymphoma (cHL), asthma, chronic obstructive pulmonary disease, eosinophilia, allergic inflammatory diseases, as well as other diseases connected with hIL-5. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-5 (SEQ ID NO:133). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 12/13, 32/33, 67/68, 76/77, 77/78, 80/81, 83/84, 84/85, 85/86, 90/91, 91/92, 92/93 and/or 98/99 (most preferred between positions 90/91, 91/92, 92/93 and/or 98/99) in hIL-5, or a peptide bond in proximity to these positions in hIL-5, or peptide bonds in protein targets related to hIL-5 at positions having structural homology or sequence homology to these positions.

In a ninth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin-6 (hIL-6). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, classical Hodgkin's Lymphoma (cHL), breast cancer, renal cell carcinoma, multiple myeloma, lymphoma, leukemia, Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, systemic lupus erythematosus, Systemic inflammatory response syndrome (SIRS) which leads to distant organ damage and multpile organ dysfunction syndrome (MODS), chronic obstructive pulmonary disease (COPD), Castleman's diseases, inflammatory bowel diseases, Crohn's disease, as well as other diseases connected with hIL-6. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-6 (SEQ ID NO:134). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 32/33, 35/36, 55/56, 71/72, 129/130, 130/131, 132/133, 135/136, 141/142, 161/162, 180/181 and/or 183/184 (most preferred between positions 135/136 and/or 141/142) in hIL-6, or a peptide bond in proximity to these positions in hIL-6, or peptide bonds in protein targets related to hIL-6 at positions having structural homology or sequence homology to these positions.

In a tenth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin 8 (hIL-8). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, Crohn's disease, Ulcerative colitis, classical Hodgkin's Lymphoma (cHL), Systemic inflammatory response syndrome (SIRS) which leads to distant organ damage and multple organ dysfunction syndrome (MODS), chronic obstructive pulmonary disease (COPD), endometriosis, psoriasis and atherosclerotic lesions, as well as other diseases connected with hIL-8. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-8 (SEQ ID NO:100). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 1/12, 15/16, 45/46, 47/48, 52/53, 54/55, 60/61, 64/65 and/or 67/68 (most preferred between positions 45/46) in hIL-8, or a peptide bond in proximity to these positions in hIL-8, or peptide bonds in protein targets related to hIL-8 at positions having structural homology or sequence homology to these positions.

In a eleventh embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin-10 (hIL-10). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, classical Hodgkin's Lymphoma (cHL) and diseases related to the suppression of cytotoxic T-cells, as well as other diseases connected with hIL-10. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-10 (SEQ ID NO:135). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 24/25, 25/26, 27/28, 28/29, 40/41, 44/45, 49/50, 57/58, 59/60, 84/85, 86/87, 106/107, 107/108, 110/111, 130/131, 134/135, 137/138, 138/139 and/or 144/145 (most preferred between positions 24/25, 27/28, 44/45, 49/50, 86/87, 137/138 and/or 144/145) in hIL-10, or a peptide bond in proximity to these positions in hIL-10, or peptide bonds in protein targets related to hIL-10 at positions having structural homology or sequence homology to these positions.

In a twelfth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin 12 beta chain (hIL-12ß). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, Crohn's disease and classical Hodgkin's Lymphoma (cHL), as well as other diseases connected with hIL-12ß. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-12ß (SEQ ID NO:97). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 14/15, 18/19, 29/30, 34/35, 87/88, 99/100, 102/103, 104/105, 161/162, 174/175, 222/223, 225/226, 228/229, 238/239, 268/269 and/or 293/294 (most preferred between positions 18/19, 34/35, 87/88 and/or 161/162) in hIL-12ß, or a peptide bond in proximity to these positions in hIL-12ß, or peptide bonds in protein targets related to hIL-12ß at positions having structural homology or sequence homology to these positions.

In a thirteenth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin 13 (hIL-13). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of cancer, such as, but not limited to, classical Hodgkin's Lymphoma (cHL), eosinophilia, asthma, chronic obstructive pulmonary disease, fibrosis, psoriasis,

atopic dermatitis and Ulcerative colitis, as well as other diseases connected with hIL-13. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-13 (SEQ ID NO:119). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 25/26, 62/63, 65/66, 86/87, 87/88, 98/99, 108/109 and/or 111/112 (most preferred between positions 87/88) in hIL-13, or a peptide bond in proximity to these positions in hIL-13, or peptide bonds in protein targets related to hIL-13 at positions having structural homology or sequence homology to these positions.

In a fourteenth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interleukin 18 (hIL-18). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, Crohn's disease, inflammation liver injuries, pulmonary tuberculosis, plural tuberculosis and rheumatoid arthritis, as well as other diseases connected with hIL-18. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIL-18 (SEQ ID NO:98). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 17/18, 32/33, 37/38, 39/40, 40/41, 53/54, 58/59, 79/80, 90/91, 93/94, 98/99, 110/111, 120/121, 123/124, 131/132, 132/133, 142/143, 147/148 and/or 157/158 (most preferred between positions 37/38, 132/133, 142/143 and/or 157/158) in hIL-18, or a peptide bond in proximity to these positions in hIL-18, or peptide bonds in protein targets related to hIL-18 at positions having structural homology or sequence homology to these positions.

In a fifteenth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interferon-gamma (hIFN-gamma). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, classical Hodgkin's Lymphoma (cHL), Crohn's disease and type I diabetes, as well as other diseases connected with hIFN-gamma. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIFN-gamma (SEQ ID NO:137). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 2/3, 6/7, 13/14, 21/22, 24/25, 34/35, 36/37, 37/38, 62/63, 68/69, 83/84, 86/87, 90/91, 102/103, 107/108 and/or 108/109 (most preferred

between positions 13/14, 24/25, 37/38, 62/63, 68/69, 102/103 and/or 107/108) in hIFN-gamma, or a peptide bond in proximity to these positions in hIFN-gamma, or peptide bonds in protein targets related to hIFN-gamma at positions having structural homology or sequence homology to these positions.

In a sixteenth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human small inducible cytokine A2 (hCCL2). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, Crohn's disease and Ulcerative colitis, as well as other diseases connected with hCCL2. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hCCL2 (SEQ ID NO:102). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 3/4, 13/14, 18/19, 19/20, 24/25, 29/30, 38/39, 54/55, 56/57, 58/59, 62/63, 65/66 and/or 68/69 (most preferred between positions 19/20, 29/30, 38/39, 54/55 and/or 62/63) in hCCL2, or a peptide bond in proximity to these positions in hCCL2, or peptide bonds in protein targets related to hCCL2 at positions having structural homology or sequence homology to these positions.

In a seventeenth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Eotaxin (hCCL11). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, Crohn's disease and Ulcerative colitis, classical Hodgkin's Lymphoma (cHL), chronic pathophysiologic dysfunction, characterized by an influx mainly of Th2 cells, and eosinophilia, as well as other diseases connected with hCCL11. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hCCL11 (SEQ ID NO:101). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 11/12, 16/17, 17/18, 22/23, 27/28, 33/34, 44/45, 47/48, 48/49, 52/53, 54/55, 56/57, 60/61, 66/67 and/or 73/74 (most preferred between positions 48/49 and/or 66/67) in hCCL11, or a peptide bond in proximity to these positions in hCCL11, or peptide bonds in protein targets related to hCCL11 at positions having structural homology or sequence homology to these positions.

In an eighteenth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Vascular endothelial growth factor (hVEGF). The

enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, all solid tumors and metastatic solid tumors, inflammatory breast cancer, as well as other diseases connected with hVEGF. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hVEGF (SEQ ID NO:103). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 16/17, 19/20, 23/24, 34/35, 41/42, 56/57, 62/63, 63/64, 64/65, 65/66, 82/83, and/or 84/85 (most preferred between positions 23/24, 41/42, 63/64, 82/83 and/or 84/85) in hVEGF, or a peptide bond in proximity to these positions in hVEGF, or peptide bonds in protein targets related to hVEGF at positions having structural homology or sequence homology to these positions.

In an ninteenth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Transforming growth factor beta 1 (hTGF-ß1). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, a variety of cancers, including breast cancer, colorectal cancer and classical Hodgkin's Lymphoma (cHL), fibrosis, suppression of cell-mediated immunity, glaucoma, diffuse systemic sclerosis as well as other diseases connected with hTGF-ß1. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hTGF-ß1 (SEQ ID NO:104). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 23/24, 25/26, 26/27, 27/28, 37/38, 55/56 and/or 94/95 (most preferred between positions 25/26, 55/56 and/or 94/95) in hTGF-ß1, or a peptide bond in proximity to these positions in hTGF-ß1, or peptide bonds in protein targets related to hTGF-ß1 at positions having structural homology or sequence homology to these positions.

In a twentieth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Somatotropin (human Growth hormone; hGH). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, acromegaly, diabetes and diabetic kidney disease including renal hypertrophy and glomerular enlargement and cardiovascular disorders, as well as other diseases connected with hGH. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hGH (SEQ ID NO:121). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 8/9,

16/17, 19/20, 26/27, 33/34, 38/39, 41/42, 70/71, 77/78, 94/95, 103/104, 112/113, 115/116, 116/117, 130/131, 147/148, 154/155 and/or 178/179 (most preferred between positions 112/113, 147/148 and/or 154/155) in hGH, or a peptide bond in proximity to these positions in hGH, or peptide bonds in protein targets related to hGH at positions having structural homology or sequence homology to these positions.

In a twenty-first embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Insulin-like growth factor II (hIGF-II). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, diabetes and diabetic kidney disease, as well as other diseases connected with hIGF-II. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIGF-II (SEQ ID NO:122). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 15/16, 23/24, 24/25, 34/35, 37/38, 38/39, 48/49 and/or 49/50 (most preferred between positions 23/24) in hIGF-II, or a peptide bond in proximity to these positions in hIGF-II, or peptide bonds in protein targets related to hIGF-II at positions having structural homology or sequence homology to these positions.

In a twenty-second embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Hepatocyte growth factor (hHGF). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, angiogenic disorders and hepatocellular carcinoma, as well as other diseases connected with hHGF. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hHGF (SEQ ID NO:120). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 54/55, 60/61, 62/63, 63/64, 68/69, 76/77, 112/113, 123/124, 134/135, 168/169, 198/199 and/or 202/203 (most preferred between positions 63/64, 68/69, 76/77, 168/169 and/or 202/203) in hHGF, or a peptide bond in proximity to these positions in hHGF, or peptide bonds in protein targets related to hHGF at positions having structural homology or sequence homology to these positions.

In a twenty-third embodiment of this variant the enzyme is a protease and is capable of hydrolysing human hInsulin (hInsulin). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of

diseases, such as, but not limited to, insulin overdosage, as well as other diseases connected with hInsulin. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hInsulin B chain (SEQ ID NO:105) and/or hInsulin A chain (SEQ ID NO:106). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 16/17 and/or 22/23 in hInsulin B and/or between position 14/15 in Insulin A, or a peptide bond in proximity to these positions in hInsulin A or B, or peptide bonds in protein targets related to hInsulin A or B at positions having structural homology or sequence homology to these positions.

In a twenty-fourth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human hGhrelin (hGhrelin). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, obesity, as well as other diseases connected with hGhrelin. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hGhrelin (SEQ ID NO:107). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 1/2, 2/3, 3/4 and/or 4/5 in hGhrelin, or a peptide bond in proximity to these positions in hGhrelin, or peptide bonds in protein targets related to hGhrelin at positions having structural homology or sequence homology to these positions.

In a twenty-fifth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human angiotensinogen (angiotensin). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, essential hypertension, as well as other diseases connected with angiotensin. Preferably, said enzyme or said fusion protein is capable of specifically inactivating angiotensin (SEQ ID NO:108). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 1/2, 3/4 and/or 7/8 (most preferred between positions 3/4) in angiotensin, or a peptide bond in proximity to these positions in angiotensin, or peptide bonds in protein targets related to angiotensin at positions having structural homology or sequence homology to these positions.

In a twenty-sixth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human leptin precursor (leptin). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of

diseases, such as, but not limited to, obesity, as well as other diseases connected with leptin. Preferably, said enzyme or said fusion protein is capable of specifically inactivating leptin (SEQ ID NO:127). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 8/9, 9/10, 15/16, 23/24, 40/41, 53/54, 71/72, 85/86, 94/95, 108/109 and/or 141/142 (most preferred between positions 9/10, 40/41, 71/72, 94/95 and/or 108/109) in leptin, or a peptide bond in proximity to these positions in leptin, or peptide bonds in protein targets related to leptin at positions having structural homology or sequence homology to these positions.

In a twenty-seventh embodiment of this variant the enzyme is a protease and is capable of hydrolysing Protective antigen (PA-83). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, anthrax infection, as well as other diseases connected with PA-83. Preferably, said enzyme or said fusion protein is capable of specifically inactivating PA-83 (SEQ ID NO:123). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 72/73, 73/74, 92/93, 93/94, 131/132, 149/150, 178/179, 213/214, 214/215, 387/388, 425/426, 426/427, 427/428, 453/454, 520/521, 608/609, 617/618, 671/672, 679/680, 680/681, 683/684 and/or 684/685 (most preferred between positions 72/73, 73/74, 93/94, 149/150, 387/388, 425/426, 427/428 and/or 683/684) in PA-83, or a peptide bond in proximity to these positions in PA-83, or peptide bonds in protein targets related to PA-83 at positions having structural homology or sequence homology to these positions.

In a twenty-eighth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human plasminogen (plasminogen). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, thrombosis, as well as other diseases connected with plasminogen. Preferably, said enzyme or said fusion protein is capable of specifically inactivating plasminogen (SEQ ID NO:140). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bond between position 580/581 in plasminogen, or a peptide bond in proximity to this position in plasminogen, or peptide bonds in protein targets related to plasminogen at positions having structural homology or sequence homology to these positions.

In a twenty-ninth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Prothrombin (thrombin). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, bleeding, as well as other diseases connected with thrombin. Preferably, said enzyme or said fusion protein is capable of specifically inactivating thrombin (SEQ ID NO:149). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 198/199, 327/328, 363/364 (most preferred between positions 327/328 and/or 363/364) in thrombin, or a peptide bond in proximity to these positions in thrombin, or peptide bonds in protein targets related to thrombin at positions having structural homology or sequence homology to these positions

In a thirty embodiment of this variant the enzyme is a protease and is capable of hydrolysing human beta-secretase. The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, Alzheimer, as well as other diseases connected with human beta-secretase precursor. Preferably, said enzyme or said fusion protein is capable of specifically inactivating human beta-secretase precursor (SEQ ID NO:139). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 61/62, 64/65, 159/160, 238/239, 239/240, 246/247, 256/257, 330/331 and/or 365/366 (most preferred between positions 61/62, 246/247 and/or 365/366) in human beta-secretase precursor, or a peptide bond in proximity to these positions in human beta-secretase precursor, or peptide bonds in protein targets related to human beta-secretase precursor at positions having structural homology or sequence homology to these positions.

In a thirty-first embodiment of this variant the enzyme is a protease and is capable of hydrolysing human matrix metalloproteinase-2 (hMMP-2). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, a variety of cancers including bladder cancer, breast tumor cancer, gastric cancer and lung cancer, as well as other diseases connected with hMMP-2. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hMMP-2 (SEQ ID NO:131). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 62/63, 68/69, 75/76, 76/77, 79/80, 88/89,

110/111, 112/113, 115/116, 120/121, 164/165, 254/255, 267/268, 296/297, 324/325, 325/326, 382/383, 383/384, 470/471, 500/501, 550/551, 564/565, 595/596, 597/598, 608/609, 646/647, 649/650 and/or 650/651 (most preferred between positions 68/69, 115/116, 120/121, 164/165, 325/326, 383/384, 470/471, 500/501, 595/596, 608/609 and/or 650/651) in hMMP-2, or a peptide bond in proximity to these positions in hMMP-2, or peptide bonds in protein targets related to hMMP-2 at positions having structural homology or sequence homology to these positions.

In a thirty-second embodiment of this variant the enzyme is a protease and is capable of hydrolysing human matrix metalloproteinase-9 (hMMP-9). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, a variety of cancers including bladder cancer, breast tumor cancer, gastric cancer and lung cancer, as well as other diseases connected with hMMP-9. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hMMP-9 (SEQ ID NO:132). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 41/42, 42/43, 106/107, 113/114, 134/135, 160/161, 162/163, 163/164, 222/223, 226/227, 265/266, 266/267, 267/268, 284/285, 309/310, 321/322, 322/323, 324/325, 356/357, 380/381, 433/434 and/or 440/441 (most preferred between positions 160/161, 163/164, 226/227, 284/285, 321/322, 322/323 and/or 433/434) in hMMP-9, or a peptide bond in proximity to these positions in hMMP-9, or peptide bonds in protein targets related to hMMP-9 at positions having structural homology or sequence homology to these positions.

In a thirty-third embodiment of this variant the enzyme is a protease and is capable of hydrolysing HIV membrane glycoprotein (GP120). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, AIDS or HIV infection, as well as other diseases connected with GP120 or HIV infection. Preferably, said enzyme or said fusion protein is capable of specifically inactivating GP120 (SEQ ID NO:124). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 97/98, 99/100, 107/108, 113/114, 117/118, 227/228, 231/233, 279/280, 335/336, 337/338, 368/369, 412/413, 419/420, 429/430, 444/445, 457/458, 474/475, 476/477, 477/478, 485/486 and/or

490/491 (most preferred between positions 99/100, 368/369, 412/413, 419/420, 444/445 and/or 490/491) in GP120, or a peptide bond in proximity to these positions in GP120, or peptide bonds in protein targets related to GP120 at positions having structural homology or sequence homology to these positions.

In a thirty-fourth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Cytotoxic T-lymphocyte protein 4 (hCTLA-4). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, breast cancer, as well as other diseases connected with hCTLA-4. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hCTLA-4 (SEQ ID NO:144). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 14/15, 28/29, 33/34, 38/39, 41/42, 62/63, 72/73, 85/86, 95/96, 100/101, 105/106, 119/120, 125/126 and/or 127/128 (most preferred between positions 14/15, 28/29, 38/39, 41/42, 62/63 and/or 85/86) in hCTLA-4, or a peptide bond in proximity to these positions in hCTLA-4, or peptide bonds in protein targets related to hCTLA-4 at positions having structural homology or sequence homology to these positions.

In a thirty-fifth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Integrin alpha-2 (hVLA-2). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, renal tumors, uveal melanomas and gastrointestinal tumors, as well as other diseases connected with hVLA-2. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hVLA-2 (SEQ ID NO:147). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 160/161, 174/175, 201/202, 219/220, 231/232, 232/233, 233/234, 243/244, 259/260, 264/265, 268/269, 288/289, 292/293, 294/295, 298/299, 301/302, 310/311 and/or 317/318 (most preferred between positions 160/161, 174/175, 201/202, 219/220, 243/244, 264/265, 292/293 and/or 294/295) in hVLA-2, or a peptide bond in proximity to these positions in hVLA-2, or peptide bonds in protein targets related to hVLA-2 at positions having structural homology or sequence homology to these positions.

In a thirty-sixth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Vascular endothelial growth factor receptor 1

(hVEGFR 1). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, solid tumors and metastatic solid tumors, astrocytic brain tumors, pancreatic cancer, metastatic renal cancer, metastatic solid tumors, as well as other diseases connected with hVEGFR 1. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hVEGFR 1 (SEQ ID NO:114). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 189/190, 190/191, 224/225 and/or 331/332 (most preferred between positions 189/190 and/or 331/332) in hVEGFR 1, or a peptide bond in proximity to these positions in hVEGFR 1, or peptide bonds in protein targets related to hVEGFR 1 at positions having structural homology or sequence homology to these positions.

In a thirty-seventh embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Vascular endothelial growth factor receptor 2 (hVEGFR 2). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, solid tumors and metatstatic solid tumors, pancreatic cancer, metastatic renal cancer, metastatic CRC, as well as other diseases connected with hVEGFR 2. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hVEGFR 2 (SEQ ID NO:115). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 214/215, and/or 323/324 (most preferred between position 214/215) in hVEGFR 2, or a peptide bond in proximity to these positions in hVEGFR 2, or peptide bonds in protein targets related to hVEGFR 2 at positions having structural homology or sequence homology to these positions.

In a thirty-eighth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Epidermal growth factor receptor (hEGFr). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of disesaes, such as, but not limited to, bladder cancer, breast cancer, cervical cancer, colorectal cancer, endometrial cancer, oesophageal cancer, head and neck cancer, gastric cancer, non-small-cell lung carcinoma and ovarian cancer, as well as other diseases connected with hEGFr. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hEGFr (SEQ ID NO:116). More preferably said enzyme or said fusion protein is capable of

hydrolysing the peptide bonds between positions 20/21, 29/30, 48/49, 74/75, 165/166, 202/203, 220/221, 246/247, 251/252, 269/270, 270/271, 304/305, 305/306, 357/358, 430/431, 443/444, 454/455, 455/456, 463/464, 465/466, 476/477, 507/508 and/or 509/510 in hEGFr, or a peptide bond in proximity to these positions in hEGFr, or peptide bonds in protein targets related to hEGFr at positions having structural homology or sequence homology to these positions.

In a thirty-ninth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Epithelial cell adhesion molecule (hEp-CAM). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, colorectal cancer, as well as other diseases connected with hEp-CAM. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hEp-CAM (SEQ ID NO:125). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 14/15, 19/20, 25/26, 30/31, 33/34, 55/56 an/or 70/71 (most preferred between positions 14/15, 30/31 and/or 70/71) in hEp-CAM, or a peptide bond in proximity to these positions in hEp-CAM, or peptide bonds in protein targets related to hEp-CAM at positions having structural homology or sequence homology to these positions.

In a forty embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Insulin-like growth factor I receptor (hIGF-1r). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, a variety of cancers including breast cancer, as well as other diseases connected with hIGF-1r. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIGF-1r (SEQ ID NO:126). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 59/60, 115/116, 146/147, 171/172, 191/192, 290/291, 306/307, 307/308, 335/336, 336/337, 455/456 and/or 470/471 (most preferred between positions 306/307, 307/308, 335/336 and/or 470/471) in hIGF-1r, or a peptide bond in proximity to these positions in hIGF-1r, or peptide bonds in protein targets related to hIGF-1r at positions having structural homology or sequence homology to these positions.

In a forty-first embodiment of this variant the enzyme is a protease and is capable of hydrolysing human T-cell surface antigen CD2 precursor (hCD2). The enzymes or the fusion protein can thus be used for preparing medicaments for

the treatment of diseases, such as, but not limited to, psoriasis, as well as other diseases connected with hCD2. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hCD2 (SEQ ID NO:128). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 42/43, 43/44, 48/49, 49/50, 51/52, 54/55, 63/64, 69/70, 89/90 and/or 91/92 (most preferred between positions 43/44, 51/52 and/or 89/90) in hCD2, or a peptide bond in proximity to these positions in hCD2, or peptide bonds in protein targets related to hCD2 at positions having structural homology or sequence homology to these positions.

In a forty-second embodiment of this variant the enzyme is a protease and is capable of hydrolysing human T-cell surface glycoprotein CD4 (hCD4). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, psoriasis, transplant rejection, graft-versus-host colitis, autoimmune disorders and rheumatoid arthritis, as well as other diseases connected with hCD4. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hCD4 (SEQ ID NO:129). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 166/167, 167/168, 206/207, 219/220, 224/225, 226/227, 251/252, 252/253, 322/323, 329/330 and/or 334/335 (most preferred between positions 206/207, 219/220, 251/252 and/or 252/253) in hCD4, or a peptide bond in proximity to these positions in hCD4, or peptide bonds in protein targets related to hCD4 at positions having structural homology or sequence homology to these positions.

In a forty-third embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Integrin alpha-L (hCD11a). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, psoriasis as well as other diseases connected with hCD11a. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hCD11a (SEQ ID NO:130). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 145/146, 152/153, 156/157, 159/160, 160/161, 177/178, 178/179, 189/190, 190/191, 191/192, 193/194, 197/198, 200/201, 221/222, 229/230, 249/250, 253/254, 268/269, 290/291, 297/298, 304/305 and/or 305/306 (most preferred between positions 145/146, 159/160, 160/161,

189/190, 229/230, 249/250, 268/269, 297/298, 304/305 and/or 305/306) in hCD11a, or a peptide bond in proximity to these positions in hCD11a, or peptide bonds in protein targets related to hCD11a at positions having structural homology or sequence homology to these positions.

In a forty-fourth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Interferon-gamma receptor alpha chain (hIFN-gamma-R1). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, classical Hodgkin's Lymphoma (cHL) and type I diabetes, as well as other diseases connected with hIFN-gamma-R1. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hIFN-gamma-R1 (SEQ ID NO:136). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 49/50, 52/53, 62/63, 106/107, 122/123, 174/175, 215/216 and/or 222/223 (most preferred between positions 49/50, 122/123, 174/175 and/or 215/216) in hIFN-gamma-R1, or a peptide bond in proximity to these positions in hIFN-gamma-R1, or peptide bonds in protein targets related to hIFN-gamma-R1 at positions having structural homology or sequence homology to these positions.

In a forty-fifth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Platelet membrane glycoprotein IIb/IIIa (hGPIIb/IIIa). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, unstable angina, carotid stenting, ischemic stroke, peripheral vascular diseases, angiogenesis-related diseases and disseminating tumors, as well as other diseases connected with hGPIIb/IIIa. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hGPIIb/IIIa (SEQ ID NO:141). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 67/68, 91/92, 129/130, 143/144, 144/145, 181/182, 208/209, 209/210, 216/217, 239/240, 261/262, 410/411, 532/533, 556/557, 557/558, 597/598, 650/651 and/or 689/690 (most preferred between positions 67/68, 261/262, 410/411, 650/651 and/or 689/690) in hGPIIb/IIIa, or a peptide bond in proximity to these positions in hGPIIb/IIIa, or peptide bonds in protein targets related to hGPIIb/IIIa at positions having structural homology or sequence homology to these positions.

In a forty-sixth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Intercellular adhesion molecule-1 (hICAM-1). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, Crohn's disease, as well as other diseases connected with hICAM-1. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hICAM-1 (SEQ ID NO:142). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 40/41, 88/89, 97/98, 102/103, 128/129, 131/132, 132/133, 149/150, 150/151, 160/161 and/or 166/167 (most preferred between positions 88/89, 102/103, 150/151, 160/161 and/or 166/167) in hICAM-1, or a peptide bond in proximity to these positions in hICAM-1, or peptide bonds in protein targets related to hICAM-1 at positions having structural homology or sequence homology to these positions.

In a forty-seventh embodiment of this variant the enzyme is a protease and is capable of hydrolysing human TGF-beta receptor type II (hTGF-beta RII). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, diffuse systemic sclerosis, as well as other diseases connected with hTGF-beta RII. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hTGF-beta RII (SEQ ID NO:145). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 32/33, 34/35, 35/36, 66/67, 67/68, 69/70, 82/83, 103/104, 104/105, 105/106, 118/119, 122/123 and/or 130/131 (most preferred between positions 32/33, 34/35, 66/67, 69/70, 104/105, 122/123 and/or 130/131) in hTGF-beta RII, or a peptide bond in proximity to these positions in hTGF-beta RII, or peptide bonds in protein targets related to hTGF-beta RII at positions having structural homology or sequence homology to these positions.

In a forty-eighth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Membrane cofactor protein (hMCP). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, renal tumors, uveal melanomas and gastrointestinal tumors, as well as other diseases connected with hMCP. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hMCP (SEQ ID NO:146). More preferably said enzyme or said fusion

protein is capable of hydrolysing the peptide bonds between positions 15/16, 17/18, 25/26, 31/32, 32/33, 35/36, 48/49, 67/68, 69/70, 110/111, 119/120 and/or 125/126 (most preferred between positions 15/16, 32/33, 48/49, 119/120 and/or 125/126) 130/131) in hMCP, or a peptide bond in proximity to these positions in hMCP, or peptide bonds in protein targets related to hMCP at positions having structural homology or sequence homology to these positions.

In a forty-ninth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Protease activated receptor 1 (hPAR1). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, thrombosis, as well as other diseases connected with hPAR1. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hPAR1 (SEQ ID NO:110). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 46/47, 51/52 and/or 52/53 in PAR1, or a peptide bond in proximity to these positions in hPAR1, or peptide bonds in protein targets related to hPAR1 at positions having structural homology or sequence homology to these positions.

In a fifth embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Protease activated receptor 2 (hPAR2). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, Crohn's disease, Ulcerative colitis and Inflammatory bowel disease, asthma, inflammation associated pain and arthritis, as well as other diseases connected with hPAR2. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hPAR2 (SEQ ID NO:111). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 41/42, 51/52 and/or 59/60 in hPAR2, or a peptide bond in proximity to these positions in hPAR2, or peptide bonds in protein targets related to hPAR2 at positions having structural homology or sequence homology to these positions.

In a fifty-first embodiment of this variant the enzyme is a protease and is capable of hydrolysing human Protease activated receptor 4 (hPAR4). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, thrombosis, as well as other diseases connected with hPAR4. Preferably, said enzyme or said fusion

protein is capable of specifically inactivating hPAR4 (SEQ ID NO:113). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 68/69, 74/75 and/or 78/79 in hPAR4, or a peptide bond in proximity to these positions in hPAR4, or peptide bonds in protein targets related to hPAR4 at positions having structural homology or sequence homology to these positions.

In a fifty-second embodiment of this variant the enzyme is a protease and is capable of hydrolysing human 5-hydroxytryptamine 1A receptor (h5-HT-1A). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, irritable bowel syndrome, as well as other diseases connected with h5-HT-1A. Preferably, said enzyme or said fusion protein is capable of specifically inactivating h5-HT-1A (SEQ ID NO:117). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 101/102, 102/103, 181/182 and/or 370/371 in h5-HT-1A a peptide bond in proximity to these positions in h5-HT-1A, or peptide bonds in protein targets related to h5-HT-1A at positions having structural homology or sequence homology to these positions.

In a fifty-third embodiment of this variant the enzyme is a protease and is capable of hydrolysing human carcinoembryonic antigen (hCEA). The enzymes or the fusion protein can thus be used for preparing medicaments for the treatment of diseases, such as, but not limited to, colon cancer, as well as other diseases connected with hCEA. Preferably, said enzyme or said fusion protein is capable of specifically inactivating hCEA (SEQ ID NO:138). More preferably said enzyme or said fusion protein is capable of hydrolysing the peptide bonds between positions 17/18, 69/70, 71/72, 74/75, 77/78, 98/99, 116/117, 126/127 and/or 128/129 in hCEA, or a peptide bond in proximity to these positions in hCEA, or peptide bonds in protein targets related to hCEA at positions having structural homology or sequence homology to these positions.

It is obvious to someone skilled in the art that also polymorphisms of all target sequences referred to are included. The expression "proximity to these positions" in all embodiments above refer to positions of peptide bonds that are between 10 and 5 Ångström and/or 5 amino acids, preferably 3 amino acids, next to the positions of the peptide bonds

Preferably, in this variant the scaffold of the engineered enzyme provided in step (c) is of human origin in order to avoid or reduce immunogenicity or allergenic effects associated with the application of the enzyme in the human body. Alternatively, immunoenicity and allergenicity can be reduced by deimmunization of the engineered enzyme.

In further embodiment of this variant, the target substrate is a pro-drug which is activated by the engineered enzyme. In a particular embodiment of this variant, the engineered enzyme has proteolytic activity and the target substrate is a protein target which is proteolytically activated. Examples of such pro-drugs are pro-proteins such as the inactivated forms of coagulations factors. In another particular variant, the engineered enzyme is an oxidoreductase and the target substrate is a chemical that can be activated by oxidation.

In a second variant of this aspect of the invention, the engineered enzyme is used for diagnostic puposes. In a particular embodiment of this variant, the engineered enzyme is target-specific protease. Such diagnostic purposes comprise but are not limited to applications with the aim of diagnosing diseases, testing genetic predispositions or monitoring disease progression during therapy. In a particular embodiment, the diagnosis is based on the testing for the presence or absence of a disease-specific marker protein or a disease-specific variant of a human protein in test samples such as human tissue samples, blood samples or other samples taken from patients. The testing employs a protease with specificity for a particular, disease-related target protein. The testing is done by analysing the proteolytic degradation of such protein in the test sample. In a preferred embodiment the aim of the diagnostic test is to detect and/or quantify a disease-specific variant of a native human protein. Such a diagnostic test employs a protease that is specific for the disease-related protein variant, i.e. it has significantly higher proteolytic activity on the disease-related protein variant compared to the native human protein. The disease-related protein variant is therefore detected and/or quantified by detecting and/or quantifying the activity of the target-specific protease. Such detection and/or quantification is done by directly measuring the degradation products of the target protein or indirectly by measuring the influence of the target protein on the activity of the

target-specific protease by a competition assay. In another preferred embodiment the aim of the diagnostic test is to detect and/or quantify a protein that is specific for an infection by an infectious agent such as a virus or a bacterium. Such a diagnostic test employs a protease that is specific for a protein specifically expressed upon infection by the infectious agent, i.e. it has significantly higher proteolytic activity on a particular infection-indicating protein compared to any other native human protein. The infection-indicating protein is therefore detected and/or quantified by detecting and/or quantifying the proteolytic activity of the target-specific protease. Such detection and/or quantification is done by directly measuring the degradation products of the infection-indicating protein or indirectly by measuring the influence of the infection-indicating protein on the activity of the target-specific protease by a competition assay.

In a third variant of this aspect of the invention, the engineered enzyme is used as a technical means in order to catalyze an industrially or nutritionally relevant reaction with defined specificity. In a particular embodiment of this variant the engineered enzyme has proteolytic activity, the catalyzed reaction is a proteolytic processing, and the engineered enzyme specifically hydrolyses one or more industrially or nutrionally relevant protein substrates. In a preferred embodiment of this variant the engineered enzyme hydrolyses one or more industrially or nutrionally relevant protein substrates at specific sites, thereby leading to industrially or nutrionally desired product properties such as texture, taste or precipitation characteristics. In a further particular embodiment of this variant, the engineered enzyme catalyzes the hydrolysis of glycosidic bonds (glycosidase or glycosylases activity). Then, preferably, the catalyzed reaction is a polysaccharide processing, and the engineered enzyme specifically hydrolyses one or more industrially, technically or nutrionally relevant polysaccharide substrates. In a further particular embodiment of this variant, the engineered enzyme catalyzes the hydrolysis of triglyceride esters or lipids (lipase activity). Then, preferably, the catalyzed reaction is a lipid processing step, and the engineered enzyme specifically hydrolyses one or more industrially, technically or nutrionally relevant lipid substrates. In a further particular variant of this embodiment, the engineered enzyme catalyzes the oxidation or reduction of substrates (oxidoreductase activity). Then, preferably, the engineered enzyme

specifically oxidizes or reduces one or more industrially, technically or nutrionally relevant chemical substrates.

A second aspect of the invention discloses engineered enzymes with defined specificities. These engineered enzymes are characterized by the following components:
(a) a protein scaffold capable of catalyzing at least one chemical reaction on a substrate, and
(b) one or more specificity determining regions (SDRs) located at sites in the protein scaffold that enable the resulting engineered protein to discriminate between ar least one target substrate and one or more different substrates, wherein the SDRs are essentially synthetic peptide sequences.

Preferably, such defined specificity of the engineered enzymes is not conferred by the protein scaffold.

In principle, the protein scaffold can have a variety of primary, secondary and tertiary structures. The primary structure, i.e. the amino acid sequence, can be an engineered sequence or can be derived from any viral, prokaryotic or eukaryotic origin. For human therapeutic use, however, the protein scaffold is preferably of mammalian origin, and more preferably, of human origin. Furthermore, the protein scaffold is capable to catalyze one or more chemical reactions and has preferably only a low specificity.

Preferably, derivatives of the protein scaffold are used that have modified amino acid sequences that confer improved characteristics for the applicability as protein scaffolds. Such improved characteristics comprise, but are not limited to, stability; expression or secretion yield; folding, in particular after combination of the protein scaffold with SDRs; increased or decreased sensitivity to regulators such as activators or inhibitors; immunogenicity; catalytic rate; kM or substrate affinity.

The engineered enzymes reveal their quantitative specificity from the synthetic peptide sequences that are combined with the protein scaffold. Therefore, the engineered peptide sequences are acting as Specificity Determining Regions or

SDRs. The number, the length and the positions of such SDRs can vary over a wide range. The number of SDRs within the scaffold is at least one, preferably more than one, more preferably between two and eleven, most preferably between two and six. The SDRs have a length between one and 50 amino acid residues, preferably a length between one and 15 amino acid residues, more preferably a length between one and six amino acid residues. Alternatively, the SDRs have a length between two and 20 amino acid residues, preferably a length between two and ten amino acid residues, more preferably a length between three and eight amino acid residues.

The inventive engineered enzymes can further be desribed as antibody-like protein molecules comprising constant and variable regions, but having a non-immunoglogulin backbone and having an active site (catalytic activity) in the constant region, whereby the substrate specificity of the active site is modulated by the variable region. Preferably, as in the immunoglobulin structure, the variable regions are loops of variable length and composition that interact with a target molecule.

In a particular variant of the invention, the engineered enzymes have hydrolase activity. In a preferred variant, the engineered enzymes have proteolytic activity. Particularly preferred protein scaffolds for this variant are unspecific proteases or are parts from unspecific proteases or are otherwise derived from unspecific proteases. The expressions "derived from" or "a derivative thereof" in this respect and in the following variants and embodiments refer to derivatives of proteins that are mutated at one or more amino acid positions and/or have a homology of at least 70%, preferably 90%, more preferably 95% and most preferably 99% to the original protein, and/or that are proteolytically processed, and/or that have an altered glycosylation pattern, and/or that are covalently linked to non-protein substances, and/or that are fused with further protein domains, and/or that have C-terminal and/or N-terminal truncations, and/or that have specific insertions, substitutions and/or deletions. Alternatively, "derived from" may refer to derivatives that are combinations or chimeras of two or more fragments from two or more proteins, each of which optionally comprises any or all of the aforementioned modifications. The tertiary structure of the protein scaffold can be of any type. Preferably, however, the tertiary structure belongs to

one of the following structural classes: class S1 (chymotrypsin fold of the serine proteases family), class S8 (subtilisin fold of the serine proteases family), class SC (carboxypeptidase fold of the serine proteases family), class A1 (pepsin A fold of the aspartic proteases), or class C14 (caspase-1 fold of the cysteine proteases). Examples of proteases that can serve as the protein scaffold of engineered proteolytic enzymes for the use as human therapeutics are or are derived from human trypsin, human thrombin, human chymotrypsin, human pepsin, human endothiapepsin, human caspases 1 to 14, and/or human furin.

The defined specificity of the engineered proteolytic enzymes is a measure of their ability to discriminate between at least one target peptide or protein substrates and one or more further peptide or protein substrates. Preferably, the defined specificity refers to the ability to discriminate peptide or protein substrates that differ in other positions than the P1 site, more preferably, the defined specificity refers to the ability to discriminate peptide or protein substrates that differ in other positions than the P1 site and the P1' site. Most preferably, the engineered proteolytic enzymes distinguish target peptid or protein substrates at as many sites as is necessary to preferentially hydrolyse the target substrate versus other proteins. As an example, a therapeutically useful engineered proteolytic enzyme applied intravenously in the human body should be sufficiently specific to discriminate between the target substrate and any other protein in the human serum. Preferably, such an engineered proteolytic enzyme recognizes and discriminates peptide substrates at three or more amino acid positions, more preferably at four or more positions, and even more preferably at five or more amino acid positions. These positions may either be adjacent or non-adjacent.

In a first embodiment , the protein scaffold has a tertiary structure or fold equal or similar to the tertiary structure or fold of the S1 structural subclass of serine proteases, i. e. the chymotrypsin fold, and/or has at least 70% identity on the amino acid level to a protein of the S1 structural subclass of serine proteases. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 18-25, 38-48, 54-63, 73-86, 122-130, 148-156, 165-171 and 194-204 in human trypsin I, and more preferably at one

or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 20-23, 41-45, 57-60, 76-83, 125-128, 150-153, 167-169 and 197-201 (numbering of amino acids according to SEQ ID NO:1). The number of SDRs to be combined with this type of protein scaffold is preferably between 1 and 10, and more preferably between 2 and 4. Preferably, the protein scaffold is equal to or is a derivative or homologue of one or more of the following proteins: chymotrypsin, granzyme, kallikrein, trypsin, mesotrypsin, neutrophil elastase, pancreatic elastase, enteropeptidase, cathepsin, thrombin, ancrod, coagulation factor IXa, coagulation factor VIIa, coagulation factor Xa, activated protein C, urokinase, tissue-type plasminogen activator, plasmin, Desmodus-type plasminogen activator. More preferably, the protein scaffold is trypsin or thrombin or is a derivative or homologue from trypsin or thrombin. For the use as a human therapeutic, the trypsin or thrombin scaffold is most preferably of human origin in order to minimize the risk of an immune response or an allergenic reaction.

Preferably, derivatives with improved characteristics derived from human trypsin I or from proteins with similar tertiary structure are used. Preferred examples of such derivatives are derived from human trypsin I (SEQ ID NO:1) and comprise one or more of the following amino acid substitutions E56G; R78W; Y131F; A146T; C183R.
It is preferred that at least one of two SDRs are inserted into human trypsin I, or a derivative thereof, between residues 42 and 43 (SDR 1) and between 123 and 124 (SDR 2), respectively (numbering of amino acids according to SEQ ID NO:1). In addition the SDR 1 has a preferred length of 6 and the SDR 2 has a preferred length of 5 amino acids, respectively. In a preferred variant of this embodiment, the SDR 1 and SDR 2 sequences comprise one of the amino acid sequences listed in table 2. Such engineered proteolytic enzymes have specificity for the target substrate B as exemplified in example IV.

In a further embodiment the protein scaffold belongs to the S8 structural subclass of serine proteases and/or has a tertiary structure similar to subtilisin E from Bacillus subtilis_and/or has at least 70% identity on the amino acid level to a protein of the S8 structural subclass of serine proteases. Preferably, the scaffold belongs to the subtilisin family or the human pro-protein convertases. It

is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 6-17, 25-29, 47-55, 59-69, 101-111, 117-125, 129-137, 139-154, 158-169, 185-195 and 204-225 in subtilisin E from Bacillus subtilis, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 59-69, 101-111, 129-137, 158-169 and 204-225 (numbering of amino acids according to SEQ ID NO:7). It is preferred that the protein scaffold is equal to or is a derivative or homologue of one or more of the following proteins: subtilisin Carlsberg; B. subtilis subtilisin E; subtilisin BPN'; B. licheniformis subtilisin; B. lentus subtilisin; Bacillus alcalophilus alkaline protease; proteinase K; kexin; human pro-protein convertase; human furin. In a preferred variant, subtilisin BPN' or one of the proteins SPC 1 to 7 is used as the protein scaffold.

In a further embodiment the protein scaffold belongs to the family of aspartic proteases and/or has a tertiary structure similar to human pepsin. Preferably, the scaffold belongs to the A1 class of proteases and/or has at least 70% identity on the amino acid level to a protein of the A1 class of proteases. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 6-18, 49-55, 74-83, 91-97, 112-120, 126-137, 159-164, 184-194, 242-247, 262-267 and 277-300 in human pepsin, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 10-15, 75-80, 114-118, 130-134, 186-191 and 280-296 (numbering of amino acids according to SEQ ID NO:11). It is preferred that the protein scaffold is equal to or is a derivative or homologue of one or more of the following proteins: pepsin, chymosin, renin, cathepsin, yapsin. Preferably, pepsin or endothiopepsin or a derivative or homologue thereof is used as the protein scaffold.

In a further embodiment the protein scaffold belongs to the cysteine protease family and/or has a tertiary structure similar to human caspase 7. Preferably the scaffold belongs to the C14 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C14 class of cysteine proteases. It is preferred that SDRs are inserted into the protein scaffold at one

or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 78-91, 144-160, 186-198, 226-243 and 271-291 in human caspase 7, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 80-86, 149-157, 190-194 and 233-238 (numbering of amino acids according to SEQ ID NO:14). It is preferred that the protein scaffold is equal to or is a derivative or homologue of one of the caspases 1 to 9.

In a further embodiment the protein scaffold belongs to the S11 class of serine proteases or has at least 70% identity on the amino acid level to a protein of the S11 class of serine proteases and/or has a tertiary structure similar to D-alanyl-D-alanine transpeptidase from Streptomyces species K15. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 67-79, 137-150, 191-206, 212-222 and 241-251 in D-alanyl-D-alanine transpeptidase from Streptomyces species K15, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 70-75, 141-147, 195-202 and 216-220 (numbering of amino acids according to SEQ ID NO:15). It is preferred that the D-alanyl-D-alanine transpeptidase from Streptomyces species K15 or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the S21 class of serine proteases or has at least 70% identity on the amino acid level to a protein of the S21 class of serine proteases and/or has a tertiary structure similar to assemblin from human cytomegalovirus. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 25-33, 64-69, 134-155, 162-169 and 217-244 in assemblin from human cytomegalovirus, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 27-31, 164-168 and 222-239 (numbering of amino acids according to SEQ ID NO:16). It is preferred that the assemblin from human cytomegalovirus or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the S26 class of serine proteases or has at least 70% identity on the amino acid level to a protein of the S26 class of serine proteases and/or has a tertiary structure similar to the signal peptidase from Escherichia coli. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 8-14, 57-68, 125-134, 239-254, 200-211 and 228-239 in signal peptidase from Escherichia coli, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 9-13, 60-67, 127-132 and 203-209 (numbering of amino acids according to SEQ ID NO:17). It is preferred that the signal peptidase from Escherichia coli or a derivative or homologue thereof is used as the scaffold.

In an further embodiment the protein scaffold belongs to the S33 class of serine proteases or has at least 70% identity on the amino acid level to a protein of the S33 class of serine proteases and/or has a tertiary structure similar to the prolyl aminopeptidase from Serratia marcescens. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 47-54, 152-160, 203-212 and 297-302 in prolyl aminopeptidase from Serratia marcescens, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 50-53, 154-158 and 206-210 (numbering of amino acids according to SEQ ID NO:18). It is preferred that the prolyl aminopeptidase from Serratia marcescens or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the S51 class of serine proteases or has at least 70% identity on the amino acid level to a protein of the S51 class of serine proteases and/or has a tertiary structure similar to aspartyl dipeptidase from Escherichia coli. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 8-16, 38-46, 85-92, 132-140, 159-170 and 205-211 in aspartyl dipeptidase from Escherichia coli, and more preferably at one or more positions from the group of

positions that correspond structurally or by amino acid sequence homology to the regions 10-14, 87-90, 134-138 and 160-165 (numbering of amino acids according to SEQ ID NO:19). It is preferred that the aspartyl dipeptidase from Escherichia coli or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the A2 class of aspartic proteases or has at least 70% identity on the amino acid level to a protein of the A2 class of aspartic proteases and/or has a tertiary structure similar to the protease from human immunodeficiency virus. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 5-12, 17-23, 27-30, 33-38 and 77-83 in protease from human immunodeficiency virus, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 7-10, 18-21, 34-37 and 79-82 (numbering of amino acids according to SEQ ID NO:20). It is preferred that the protease from human immunodeficiency virus, preferably HIV-1 protease, or a derivative or homologue thereof is used as the scaffold.

In an further embodiment the protein scaffold belongs to the A26 class of aspartic proteases or has at least 70% identity on the amino acid level to a protein of the A26 class of aspartic proteases and/or has a tertiary structure similar to the omptin from Escherichia coli. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 28-40, 86-98, 150-168, 213-219 and 267-278 in omptin from Escherichia coli, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 33-38, 161-168 and 273-277 (numbering of amino acids according to SEQ ID NO:21). It is preferred that the omptin from Escherichia coli or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the C1 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C1 class of cysteine proteases and/or has a tertiary structure similar to the

papain from Carica papaya. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 17-24, 61-68, 88-95, 135-142, 153-158 and 176-184 in papain from Carica papaya, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 63-66, 136-139 and 177-181 (numbering of amino acids according to SEQ ID NO:22). It is preferred that the papain from Carica papaya or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the C2 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C2 class of cysteine proteases and/or has a tertiary structure similar to human calpain-2. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 90-103, 160-172, 193-199, 243-260, 286-294 and 316-322 in human calpain-2, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 92-101, 245-250 and 287-291 (numbering of amino acids according to SEQ ID NO:23). It is preferred that the human calpain-2 or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the C4 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C4 class of cysteine proteases and/or has a tertiary structure similar to NIa protease from tobacco etch virus. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 23-31, 112-120, 144-150, 168-176 and 205-218 in NIa protease from tobacco etch virus, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 145-149, 169-174 and 212-218 (numbering of amino acids according to SEQ ID NO:24). It is preferred that the NIa protease from tobacco etch virus (TEV protease) or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the C10 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C10 class of cysteine proteases and/or has a tertiary structure similar to the streptopain from Streptococcus pyogenes. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 81-90, 133-140, 150-164, 191-199, 219-229, 246-256, 306-312 and 330-337 in streptopain from Streptococcus pyogenes, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 82-87, 134-138, 250-254 and 331-335 (numbering of amino acids according to SEQ ID NO:25). It is preferred that the streptopain from Streptococcus pyogenes or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the C19 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C19 class of cysteine proteases and/or has a tertiary structure similar to human ubiquitin specific protease 7. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 3-15, 63-70, 80-86, 248-256, 272-283 and 292-304 in human ubiquitin specific protease 7, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 10-15, 251-255, 277-281 and 298-304 (numbering of amino acids according to SEQ ID NO:26). It is preferred that the human ubiquitin specific protease 7 or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the C47 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C47 class of cysteine proteases and/or has a tertiary structure similar to the staphopain from Staphylococcus aureus. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 15-23, 57-66, 108-119, 142-149 and 157-164 in staphopain from Staphylococcus aureus, and more preferably at one or more positions from the group of positions

that correspond structurally or by amino acid sequence homology to the regions 17-22, 111-117, 143-147 and 159-163 (numbering of amino acids according to SEQ ID NO:27). It is preferred that the staphopain from Staphylococcus aureus or a derivative or homologue thereof is used as the scaffold.

In an further embodiment the protein scaffold belongs to the C48 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C48 class of cysteine proteases and/or has a tertiary structure similar to the Ulp1 endopeptidase from Saccharomyces cerevisiae. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 40-51, 108-115, 132-141, 173-179 and 597-605 in Ulp1 endopeptidase from Saccharomyces cerevisiae, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 43-49, 110-113, 133-137 and 175-178 (numbering of amino acids according to SEQ ID NO:28). It is preferred that the Ulp1 endopeptidase from Saccharomyces cerevisiae or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the C56 class of cysteine proteases or has at least 70% identity on the amino acid level to a protein of the C56 class of cysteine proteases and/or has a tertiary structure similar to the Pfpl endopeptidase from Pyrococcus horikoshii. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 8-16, 40-47, 66-73, 118-125 and 147-153 in Pfpl endopeptidase from Pyrococcus horikoshii, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 9-14, 68-71, 120-123 and 148-151 (numbering of amino acids according to SEQ ID NO:29). It is preferred that the Pfpl endopeptidase from Pyrococcus horikoshii or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the M4 class of metallo proteases or has at least 70% identity on the amino acid level to a protein of the M4 class of metallo proteases and/or has a tertiary structure similar to

thermolysin from Bacillus thermoproteolyticus. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 106-118, 125-130, 152-160, 197-204, 210-213 and 221-229 in thermolysin from Bacillus thermoproteolyticus, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 108-115, 126-129, 199-203 and 223-227 (numbering of amino acids according to SEQ ID NO:30). It is preferred that the thermolysin from Bacillus thermoproteolyticus or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the M10 class of metallo proteases or has at least 70% identity on the amino acid level to a protein of the M10 class of metallo proteases and/or has a tertiary structure similar to human collagenase. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 2-7, 68-79, 85-90, 107-111 and 135-141 in human collagenase, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 3-6, 71-78 and 136-140 (numbering of amino acids according to SEQ ID NO:31). It is preferred that human collagenase or a derivative or homologue thereof is used as the scaffold.

It is further preferred that the engineered enzymes have glycosidase activity. A particularly suited protein scaffold for this variant is a glycosylase or is derived from a glycosylase. Preferably, the tertiary structure belongs to one of the following structural classes: class GH13, GH7, GH12, GH11, GH10, GH28, GH26, and GH18 (beta/alpha)8 barrel.

In a first embodiment the protein scaffold belongs to the GH13 class of glycosylases or has at least 70% identity on the amino acid level to a protein of the GH13 class of glycosylases and/or has a tertiary structure similar to human pancreatic alpha-amylase. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 50-60, 100-110,

148-167, 235-244, 302-310 and 346-359 in human pancreatic alpha-amylase, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 51-58, 148-155 and 303-309 (numbering of amino acids according to SEQ ID NO:32). It is preferred that human pancreatic alpha-amylase or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the GH7 class of glycosylases or has at least 70% identity on the amino acid level to a protein of the GH7 class of glycosylases and/or has a tertiary structure similar to cellulase from Trichoderma reesei. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 47-56, 93-104, 173-182, 215-223, 229-236 and 322-334 in cellulase from Trichoderma reesei, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 175-180, 218-222 and 324-332 (numbering of amino acids according to SEQ ID NO:33). It is preferred that cellulase from Trichoderma reesei or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the GH12 class of glycosylases or has at least 70% identity on the amino acid level to a protein of the GH12 class of glycosylases and/or has a tertiary structure similar to cellulase from Aspergillus niger. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 18-28, 55-60, 106-113, 126-132 and 149-159 in cellulase from Aspergillus niger, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 20-26, 56-59, 108-112 and 151-156 (numbering of amino acids according to SEQ ID NO:34). It is preferred that cellulase from Aspergillus niger or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the GH11 class of glycosylases or has at least 70% identity on the amino acid level to a protein of

the GH11 class of glycosylases and/or has a tertiary structure similar to xylanase from Aspergillus niger. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 7-14, 33-39, 88-97, 114-126 and 158-167 in xylanase from Aspergillus niger, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 20-26, 56-59, 108-112 and 151-156 (numbering of amino acids according to SEQ ID NO:35). It is preferred that xylanase from Aspergillus niger or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the GH10 class of glycosylases or has at least 70% identity on the amino acid level to a protein of the GH10 class of glycosylases and/or has a tertiary structure similar to xylanase from Streptomyces lividans. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 21-29, 42-50, 84-92, 130-136, 206-217 and 269-278 in xylanase from Streptomyces lividans, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 43-49, 86-90, 208-213 and 271-276 (numbering of amino acids according to SEQ ID NO:36). It is preferred that xylanase from Streptomyces lividans or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the GH28 class of glycosylases or has at least 70% identity on the amino acid level to a protein of the GH28 class of glycosylases and/or has a tertiary structure similar to pectinase from Aspergillus niger. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 82-88, 118-126, 171-178, 228-236, 256-264 and 289-299 in pectinase from Aspergillus niger, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 116-124, 174-178 and 291-296 (numbering of amino acids according

to SEQ ID NO:37). It is preferred that pectinase from Aspergillus niger or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the GH26 class of glycosylases or has at least 70% identity on the amino acid level to a protein of the GH26 class of glycosylases and/or has a tertiary structure similar to mannanase from Pseudomonas cellulosa. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 75-83, 113-125, 174-182, 217-224, 247-254, 324-332 and 325-340 in mannanase from Pseudomonas cellulosa, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 115-123, 176-180, 286-291 and 328-337 (numbering of amino acids according to SEQ ID NO:38). It is preferred that mannanase from Pseudomonas cellulosa or a derivative or homologue thereof is used as the scaffold.

In an further embodiment the protein scaffold belongs to the GH18 (beta/alpha)8 barrel class of glycosylases or has at least 70% identity on the amino acid level to a protein of the GH18 class of glycosylases and/or has a tertiary structure similar to chitinase from Bacillus circulans. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 21-29, 57-65, 130-136, 176-183, 221-229, 249-257 and 327-337 in chitinase from Bacillus circulans, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 59-63, 178-181, 250-254 and 330-336 (numbering of amino acids according to SEQ ID NO:39). It is preferred that chitinase from Bacillus circulans or a derivative or homologue thereof is used as the scaffold.

It is further preferred that the engineered enzymes have esterhydrolase activity. Preferably, the protein scaffold for this variant have lipase, phosphatase, phytase, or phosphodiesterase activity.

In a first embodiment the protein scaffold belongs to the GX class of esterases or has at least 70% identity on the amino acid level to a protein of the GX class of esterases and/or has a tertiary structure similar to the structure of the lipase B from Candida antarctica. Preferably, the scaffold has lipase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 139-148, 188-195, 216-224, 256-266, 272-287 in lipase B from Candida antarctica, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 141-146, 218-222, 259-263 and 275-283 (numbering of amino acids according to SEQ ID NO:40). It is preferred that lipase B from Candida antarctica or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the GX class of esterases or has at least 70% identity on the amino acid level to a protein of the GX class of esterases and/or has a tertiary structure similar to the pancreatic lipase from guinea pig. Preferably, the scaffold has lipase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 78-90, 91-100, 112-120, 179-186, 207-218, 238-247 and 248-260 in pancreatic lipase from guinea pig, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 80-87, 114-118, 209-215 and 239-246 (numbering of amino acids according to SEQ ID NO:41). It is preferred that pancreatic lipase from guinea pig or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold has a tertiary structure similar to the structure of the alkaline phosphatase from Escherichia coli or has at least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the alkaline phosphatase from Escherichia coli. Preferably, the scaffold has phosphatase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the

regions 110-122, 187-142, 170-175, 186-193, 280-287 and 425-435 in alkaline phosphatase from Escherichia coli, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 171-174, 187-191, 282-286 and 426-433 (numbering of amino acids according to SEQ ID NO:42). It is preferred that alkaline phosphatase from Escherichia coli or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold has a tertiary structure similar to the structure of the bovine pancreatic desoxyribonuclease I or has at least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the bovine pancreatic desoxyribonuclease I. Preferably, the scaffold has phosphodiesterase activity. More preferably, a nuclease, and most preferably, an unspecific endonuclease or a derivative thereof is used as the scaffold. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 14-21, 41-47, 72-77, 97-111, 135-143, 171-178, 202-209 and 242-251 in bovine pancreatic desoxyribonuclease I, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 16-19, 42-46, 136-141 and 172-176 (numbering of amino acids according to SEQ ID NO:43). It is preferred that bovine pancreatic desoxyribonuclease I or human desoxyribonuclease I or a derivative or homologue thereof is used as the scaffold.

It is further preferred that the engineered enzyme has transferase activity. A particularly suited protein scaffold for this variant is a glycosyl-, a phospho- or a methyltransferase, or is a derivative thereof. Particularly preferred protein scaffolds for this variant are glycosyltransferases or are derived from glycosyltransferases. The tertiary structure of the protein scaffold can be of any type. Preferably, however, the tertiary structure belongs to one of the following structural classes: GH13 and GT1.

In a first embodiment the protein scaffold belongs to the GH13 class of transferases or has at least 70% identity on the amino acid level to a protein of

the GH13 class of transferases and/or has a tertiary structure similar to the structure of the cyclomaltodextrin glucanotransferase from Bacillus circulans. Preferably, the scaffold has transferase activity, and more preferably a glycosyltransferase is used as the scaffold. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 38-48, 85-94, 142-154, 178-186, 259-266, 331-340 and 367-377 in cyclomaltodextrin glucanotransferase from Bacillus circulans, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 87-92, 180-185, 261-264 and 269-275 (numbering of amino acids according to SEQ ID NO:44). It is preferred that cyclomaltodextrin glucanotransferase from Bacillus circulans or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold belongs to the GT1 class of tranferases or has at least 70% identity on the amino acid level to a protein of the GT1 class of transferases and/or has a tertiary structure similar to the structure of the glycosyltransferase from Amycolatopsis orientalis A82846. Preferably the scaffold has transferase activity, and more preferably glycosyltransferase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 58-74, 130-138, 185-193, 228-236 and 314-323 in glycosyltransferase from Amycolatopsis orientalis A82846, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 61-71, 230-234 and 316-321 (numbering of amino acids according to SEQ ID NO:45). It is preferred that the glycosyltransferase from Amycolatopsis orientalis A82846 or a derivative or homologue thereof is used as the scaffold.

It is further preferred that the engineered enzymes have oxidoreductase activity. A particularly suited protein scaffold for this variant is a monooxygenase, a dioxygenase or a alcohol dehydrogenase, or a derivative thereof. The tertiary structure of the protein scaffold can be of any type.

In a first embodiment the protein scaffold has a tertiary structure similar to the structure of the 2,3-diphydroxybiphenyl dioxygenase from Pseudomonas sp. or has at least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the 2,3-diphydroxybiphenyl dioxygenase from Pseudomonas sp. Preferably, the scaffold has dioxygenase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 172-185, 198-206, 231-237, 250-259 and 282-287 in 2,3-diphydroxybiphenyl dioxygenase from Pseudomonas sp., and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 175-182, 200-204, 252-257 and 284-287 (numbering of amino acids according to SEQ ID NO:46). It is preferred that the 2,3-diphydroxybiphenyl dioxygenase from Pseudomonas sp or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold has a tertiary structure similar to the structure of the catechol dioxygenase from Acinetobacter sp. or has at least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the catechol dioxygenase from Acinetobacter sp.. Preferably, the scaffold has dioxygenase activity, and more preferably catechol dioxygenase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 66-72, 105-112, 156-171 and 198-207 in catechol dioxygenase from Acinetobacter sp., and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 107-110, 161-171 and 201-205 (numbering of amino acids according to SEQ ID NO:47). It is preferred that the catechol dioxygenase from Acinetobacter sp or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold has a tertiary structure similar to the structure of the camphor-5-monooxygenase from Pseudomonas putida or has at least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the camphor-5-monooxygenase from

Pseudomonas putida. Preferably, the scaffold has monooxygenase activity, and more preferably camphor monooxygenase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 26-31, 57-63, 84-98, 182-191, 242-256, 292-299 and 392-399 in camphor-5-monooxygenase from Pseudomonas putida, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 85-96, 183-188, 244-253, 293-298 and 393-398 (numbering of amino acids according to SEQ ID NO:48). It is preferred that the camphor-5-monooxygenase from Pseudomonas putida or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold has a tertiary structure similar to the structure of the alcohol dehydrogenase from Equus callabus or has at least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the alcohol dehydrogenase from Equus callabus. Preferably, the scaffold has alcohol dehydrogenase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 49-63, 111-112, 294-301 and 361-369 in alcohol dehydrogenase from Equus callabus, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 51-61 and 295-299 (numbering of amino acids according to SEQ ID NO:49). It is preferred that the alcohol dehydrogenase from Equus callabus or a derivative or homologue thereof is used as the scaffold.

It is further preferred that the engineered enzymes have lyase activity. A particularly suited protein scaffold for this variant is a oxoacid lyase or is a derivative thereof. Particularly preferred protein scaffolds for this variant are aldolases or synthases, or are derived thereof. The tertiary structure of the protein scaffold can be of any type, but a (beta/alpha)8 barrel structure is preferred.

In a first embodiment the protein scaffold has a tertiary structure similar to the structure of the N-acetyl-d-neuramic acid aldolase from Escherichia coli or has at

least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the N-acetyl-d-neuramic acid aldolase from Escherichia coli. Preferably, the scaffold has aldolase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 45-55, 78-87, 105-113, 137-146, 164-171, 187-193, 205-210, 244-255 and 269-276 in N-acetyl-d-neuramic acid aldolase from Escherichia coli, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 45-52, 138-144, 189-192, 247-253 and 271-275 (numbering of amino acids according to SEQ ID NO:50). It is preferred that the N-acetyl-d-neuramic acid aldolase from Escherichia coli or a derivative or homologue thereof is used as the scaffold.

In a further embodiment the protein scaffold has a tertiary structure similar to the structure of the tryptophan synthase from Salmonella typhimurium or has at least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the tryptophan synthase from Salmonella typhimurium. Preferably, the scaffold has synthase activity. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 56-63, 127-134, 154-161, 175-193, 209-216 and 230-240 in tryptophan synthase from Salmonella typhimurium, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 57-62, 155-160, 178-190 and 210-215 (numbering of amino acids according to SEQ ID NO:51). It is preferred that the tryptophan synthase from Salmonella typhimurium or a derivative or homologue thereof is used as the scaffold.

It is further preferred that the engineered enzymes have isomerase activity. A particularly suited protein scaffold for this variant is a converting aldose or a converting ketose, or is a derivative thereof.

In a first embodiment, the protein scaffold has a tertiary structure similar to the structure of the xylose isomerase from Actinoplanes missouriensis or has at least

70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the xylose isomerase from Actinoplanes missouriensis. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 18-31, 92-103, 136-147, 178-188 and 250-257 in xylose isomerase from Actinoplanes missouriensis, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 20-27, 92-99 and 180-186 (numbering of amino acids according to SEQ ID NO:52). It is preferred that the xylose isomerase from Actinoplanes missouriensis or a derivative or homologue thereof is used as the scaffold.

It is further preferred that the engineered enzymes have ligase activity. A particularly suited protein scaffold for this variant is a DNA ligase, or is a derivative thereof.

In a first embodiment, the protein scaffold has a tertiary structure similar to the structure of the DNA ligase from Bacteriophage T7 or has at least 70% identity on the amino acid level to a protein that has a tertiary structure similar to the structure of the DNA-ligase from Bacteriophage T7. It is preferred that SDRs are inserted into the protein scaffold at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 52-60, 94-108, 119-131, 241-248, 255-263 and 302-318 in DNA ligase from Bacteriophage T7, and more preferably at one or more positions from the group of positions that correspond structurally or by amino acid sequence homology to the regions 96-106, 121-129, 256-262 and 304-316 (numbering of amino acids according to SEQ ID NO:53). It is preferred that the DNA ligase from Bacteriophage T7 or a derivative or homologue thereof is used as the scaffold.

A third aspect of the invention is directed to a method for generating engineered enzymes with specificities that are qualitatively and/or quantitatively novel in combination with the protein scaffold. The inventive method comprises at least the following steps:

(a) providing a protein scaffold capable to catalyze at least one chemical reaction on at least one target substrate,

(b) generating a library of engineered enzymes or isolated engineered enzymes by combining the protein scaffold from step (a) with one or more fully or partially random peptide sequences at sites in the protein scaffold that enable the resulting engineered enzyme to discriminate between at least one target substrate and one or more different substrates and

(c) selecting out of the library of engineered enzymes generated in step (b) one or more enzymes that have defined specificities towards at least one target substrate.

In a first variant of this aspect of the invention, the inventive method comprises at least the following steps:

(a) providing a protein scaffold capable to catalyze at least one chemical reaction on at least one target substrate,

(b) generating a library of engineered enzymes or isolated engineered enzymes by inserting into the protein scaffold from step (a) one or more fully or partially random peptide sequences at sites in the protein scaffold that enable the resulting engineered enzyme to discriminate between at least one target substrate and one or more different substrates and

(c) selecting out of the library of engineered enzymes generated in step (b) one or more enzymes that have defined specificities towards at least one target substrate.

Preferably, the positions at which the one or more fully or partially random peptide sequences are combined with or inserted into the protein scaffold are identified prior to the combination or insertion.

The number of insertions or other combinations of fully or partially random peptide sequences as well as their length may vary over a wide range. The number is at least one, preferably more than one, more preferably between two and eleven, most preferably between two and six. The length of such fully or partially random peptide sequences is usually less than 50 amino acid residues. Preferably, the length is between one and 15 amino acid residues, more preferably between one and six amino acid residues. Alternatively, the length is between two and 20 amino acid residues, preferably between two and ten amino acid residues, more preferably between three and eight amino acid residues.

Preferably such insertions or other combinations are performed on the DNA level, using polynucleotides encoding such protein scaffolds and polynucleotides or oligonucleotides encoding such fully or partially random peptide sequences.

Optionally, steps (a) to (c) are repeated cyclically, whereby enzymes selected in step (c) serve as the protein scaffold in step (a) of a further cycle, and randomized peptide sequences are either inserted or, alternatively, substituted for peptide sequences that have been inserted in former cycles. Thereby, the number of inserted peptide sequences is either constant or increases over the cycles. The cycles are repeated until one or more enzymes with the intended specificities are generated.

Moreover, during or after one or more rounds of steps (a) to (c), the scaffold may be mutated at one or more positions in order to make the scaffold more acceptable for the combination with SDR sequences, and/or to increase catalytic activity at a specific pH and temperature, and/or to change the glycosylation pattern, and/or to decrease sensitivity towards enzyme inhibitors, and/or to change enzyme stability.

In a second variant of this aspect of the invention, the inventive method comprises at least the following steps:
(a) providing a first protein scaffold fragment,
(b) connecting said protein scaffold fragment via a peptide linkage with a first SDR, and optionally
(c) connecting the product of step (b) via a peptide linkage with a further SDR peptide or with a further protein scaffold fragment, and optionally
(d) repeating step (c) for as many cycles as necessary in order to generate a sufficiently specific enzyme, and
(e) selecting out of the population generated in steps (a) – (d) one or more enzymes that have the desired specificities toward the one or more target substrates.

Protein scaffold fragment means a part of the sequence of a protein scaffold. A protein scaffold is comprised of at least two protein scaffold fragments.

In a third variant of this aspect of the invention, the protein scaffold, the SDRs and the engineered enzyme are encoded by a DNA sequence and an expression system is used in order to produce the protein. In an alternative variant, the protein scaffold, the SDRs and/or the engineered enzyme are chemically synthesized from peptide building blocks.

In a fourth variant of this aspect of the invention, the inventive method comprises at least the following steps:
(a) providing a polynucleotide encoding a protein scaffold capable of catalyzing one or more chemical reactions on one or more target substrates;
(b) combining one or more fully or partially random oligonucleotide sequence with the polynucleotide encoding the protein scaffold, the fully or partially random oligonucleotide sequences being located at sites in the polynucleotide that enable the encoded engineered enzyme to discriminate between the one or more target substrates and one or more other substrates; and
(c) selecting out of the population generated in step (b) one or more polynucleotides that encode enzymes that have the defined specificities toward the one or more target substrates.

Any enzyme can serve as the protein scaffold in step (a). It can be a naturally occurring enzyme, a variant or a truncated derivate therefore, or an engineered enzyme. For human therapeutic use, the protein scaffold is preferably a mammalian enzyme, and more preferably a human enzyme. In that aspect, the invention is directed to a method for the generation of essentially mammalian, especially of essentially human enzymes with specificities that are different from specificities of any enzyme encoded in mammalian genomes or in the human genome, respectively.

According to the invention, the protein scaffold provided in step (a) of this aspect requires to be capable of catalyzing one or more chemical reactions on a target substrate. Therefore, a protein scaffold is selected from the group of potential protein scaffolds by its activity on the target substrate.

In a preferred variant of this aspect of the invention, a protein scaffold with hydrolase activity is used. Preferably, a protein scaffold with proteolytic activity

is used, and more preferably, a protease with very low specificity having basic activity on the target substrate is used as the protein scaffold. Examples of proteases from different structural classes with low substrate specificity are Papain, Trypsin, Chymotrypsin, Subtilisin, SET (trypsin-like serine protease from Streptomyces erythraeus), Elastase, Cathepsin G or Chymase. Before being employed as the protein scaffold, the amino acid sequence of the protease may be modified in order to change protein properties other than specificity, e.g catalytic activity, stability, inhibitor sensitivity, or expression yield, essentially as described in WO 92/18645, or in order to change specificity, essentially as described in EP 02020576.3 and PCT/EP03/04864.

Another option for a feasible protein scaffold are lipases. Hepatic lipase, lipoprotein lipase and pancreatic lipase belong to the "lipoprotein lipase superfamily", which in turn is an example of the GX-class of lipases (M. Fischer, J. Pleiss (2003), Nucl. Acid. Res., 31, 319-321). The substrate specificity of lipases can be characterized by their relative activity towards triglycerol esters of fatty acids and phospholipids, bearing a charged head group. Alternatively, other hydrolases such as esterases, glycosylases, amidases, or nitrilases may be used as scaffolds.

Transferases are also feasible protein scaffolds. Glycoslytransferases are involved in many biological synthesis involving a variety of donors and acceptors. Alternatively, the protein scaffold may have ligase, lyase, oxidoreductase, or isomerase activity.

In a first embodiment, the one or more fully or partially random peptide sequences are inserted at specific sites in the protein scaffold. These insertion sites are characterized by the fact that the inserted peptide sequences can act as discriminators between different substrates, i.e. as Specificity Determining Regions or SDRs. Such insertion sites can be identified by several approaches. Preferably, insertion sites are identified by analysis of the three-dimensional structure of the protein scaffolds, by comparative analysis of the primary sequences of the protein scaffold with other enzymes having different quantitative specificities, or experimentally by techniques such as alanine scanning, random mutagenesis, or random deletion, or by any combination thereof.

A first approach to identify insertion sites for SDRs bases on the three-dimensional structure of the protein scaffold as it can be obtained by x-ray crystallography or by nuclear magnetic resonance studies. Structural alignment of the protein scaffold in comparison with other enzymes of the same structural class but having different quantitative specificities reveals regions of high structural similarity and regions with low structural similarity. Such an analysis can for example be done using public software such as Swiss PDB viewer (Guex, N. and Peitsch, M.C. (1997) *Electrophoresis* 18, 2714-2723). Regions of low structural similarity are preferred SDR insertion sites.

In a second approach to identify insertion sites for SDRs, three-dimensional structures of the scaffold protein in complex with competitive inhibitors or substrate analogs are analysed. It is assumed that the binding site of a competitive inhibitor significantly overlaps with the binding site of the substrate. In that case, atoms of the protein that are within a certain distance of atoms of the inhibitor are likely to be in a similar distance to the substrate as well. Choosing a short distance, e.g. < 5 Å, will result in an ensemble of protein atoms that are in close contact with the substrate. These residues would constitute the first shell contacts and are therefore preferred insertion sites for SDRs. Once first shell contacts have been identified, second shell contacts can be found by repeating the distance analysis starting from first shell atoms. In yet another alternative of the invention the distance analysis described above is performed starting from the active site residues.

In third approach to identify insertion sites for SDRs, the primary sequence of the scaffold protein is aligned with other enzymes of the same structural class but having different quantitative specificities using an alignment algorithm. Examples of such alignment algorithms are published (Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) J. Mol. Biol. 215:403-410; "Statistical methods in Bioinformatics: an introduction" by Ewens, W. & Grant, G.R. 2001, Springer, New York). Such an alignment may reveal conserved and non-conserved regions with varying sequence homology, and, in particular, additional sequence elements in one or more enzymes compared to the scaffold protein. Conserved regions of are more likely to contribute to phenotypes shared

among the different proteins, e.g. stabilizing the three-dimensional fold. Non-conserved regions and, in particular, additional sequences in enzymes with quantitatively higher specificity (Turner, R. et al. (2002) *J. Biol. Chem.*, 277, 33068-33074) are preferred insertion sites for SDRs.

For proteases currently five families are known, namely aspartic-, cysteine-, serine-, metallo- and threonine proteases. Each family includes groups of proteases that share a similar fold. Crystallographic structures of members of these groups have been solved and are accessible through public databases, e.g. the Brookhaven protein database (H.M. Berman et al. *Nucleic Acids Research*, 28 pp. 235-242 (2000)). Such databases also include structural homologs in other enzyme classes and nonenzymatically active proteins of each class. Several tools are available to search public databases for structural homologues: SCOP - a structural classification of proteins database for the investigation of sequences and structures. (Murzin A. G. et al. (1995) *J. Mol. Biol.* 247, 536-540); CATH - Class, Architecture, Topology and Homologous superfamily: a hierarchical classification of protein domain structures (Orengo et al. (1997) *Structure* 5(8) 1093-1108); FSSP - Fold classification based on structure-structure alignment of proteins (Holm and Sander (1998) *Nucl. Acids Res.* 26 316-319); or VAST - Vector alignment search tool (Gibrat, Madej and Bryant (1996) *Current Opinion in Structural Biology* 6, 377-385).

In the above described approaches, members of structural classes are compared in order to identify insertion sites for SDRs.

In a preferred variant of these approaches serine proteases of the structural class S1 are compared with each other. Trypsin represents a member with low substrate specificity, as it requires only an arginine or lysine residue at the P_1 position. On the other hand, thrombin, tissue-type plasminogen activator or enterokinase all have a high specificity towards their substrate sequences, i.e. (L/I/V/F)XPR^ NA, CPGR^ VVGG and DDDK^ , respectively (Perona, J. & Craik, C. (1997) *J. Biol. Chem.*, 272, 29987-29990; Perona, J. & Craik, C (1995) *Protein Science*, 4, 337-360). An alignment of the amino acid sequences of these proteases is described in example 1 (Figure 2) along with the identification of SDRs.

A further example within the family of serine proteases is given by members of the structural class S8 (subtilisin fold). Subtilisin is the type protease for this class and represents an unspecific protease (Ottesen,M. & Svendsen,A. (1998) *Methods Enzymol.* 19, 199-215). Furin, PC1 and PC5 are proteases of the same structural class involved in the processing of propeptides and have a high substrate specificity (Seidah, N. & Chretien, M. (1997) Curr. Opin. Biotech., 8: 602-607; Bergeron, F. et al. (2000) *J. Mol. Endocrin.*, 24:1-22). In a preferred variant of the approach alignments of the primary amino acids sequences (Figure 4) are used to identify eleven sequence stretches longer than three amino acids which specific proteases have in addition compared to subtilisin and are therefore potential specificity determining regions. In a further variant of the approach information from the three-dimensional structure of subtilisin can be used in order to further narrow down the selection (Figure 3). Out of the eleven inserted sequence stretches, three are especially close to the active site residues, namely stretch number 7, 8 and 11 which are insertions in PC5, PC1 and all three specific proteases, respectively (Figure 3). In a preferred variant, one or several amino acid stretches of variable length and composition can be inserted into the subtilisin sequence at one or several of the eleven positions. In a more preferred variant of the approach the insertion is performed at regions 7, 8 or 11 or any combination thereof. In another preferred variant of the approach protease scaffolds other than subtilisin from the structural class S8 are used.

In a further preferred variant of this approach, aspartic acid proteases of the structural class A1 are analyzed (Rawlings, N.D. & Barrett, A.J. (1995). *Methods Enzymol.* 248, 105-120; Chitpinityol, S. & Crabbe, MJ. (1998), *Food Chemistry*, 61, 395-418). Examples for the A1 structural class of aspartic proteases are pepsin with a low as well as beta-secretase (Grüninger-Leitch, F., et al. (2002) *J. Biol. Chem.* 277, 4687-4693) and renin (Wang, W. & Liang, TC. (1994) *Biochemistry*, 33, 14636-14641) with relatively high substrate specificities. Retroviral proteases also belong to this class, although the active enzyme is a dimer of two identical subunits. The viral proteases are essential for the correct processing of the polyprotein precursor to generate functional proteins which requires a high substrate specificity in each case (Wu, J. et al. (1998) *Biochemistry*, 37, 4518-4526; Pettit, S. et al. (1991) J. Biol. Chem., 266, 14539-

14547). Pepsin is the type protease for this class and represents an unspecific protease (Kageyama, T. (2002) *Cell. Mol. Life Sci.* 59, 288-306). B-secretase and Cathepsin D (Aguilar, C. F. et al. (1995) *Adv. Exp. Med. Biol.* 362, 155-166) are proteases of the same structural class and have a high substrate specificity. In a preferred variant of the approach alignments of the primary amino acids sequences (Figure 6) are used to identify six sequence stretches longer than three amino acids which are inserted in the specific proteases compared to pepsin and are therefore potential specificity determining regions. In a further variant of the approach information from the three-dimensional structure of b-secretase can be used in order to further narrow down the selection. Out of the six inserted sequence stretches, three are especially close to the active site residues, namely stretch number 1, 3 and 4 which are insertions in cathepsin D and beta-secretase, respectively (Figure 5). In a preferred variant of the approach, one or several amino acid stretches of variable length and composition can be inserted into the pepsin sequence at one or several of the six positions. In a more preferred embodiment of the invention the insertion is performed at the positions 1, 3 or 4 or any combination thereof. In another preferred embodiment of the invention protease scaffolds other than pepsin are used.

There are cases where a certain structural class does not include known members of bw and high specificity. This is exemplified by the C14 class of caspases which belong to the cysteine protease family (Rawlings, N.D. & Barrett, A.J. (1994) *Methods Enzymol.* 244, 461-486) and which all show high specificity for P_4 to P_1 positions. For example, caspase-1, caspase-3 and caspase-9 recognize the sequences YVAD^ , DEVD^ or LEHD^ , respectively. Identification of the regions that differ between the caspases will include the regions responsible for the differences in substrate specificity (Figures 7 and 8).

Finally, non-enzymatic proteins of the same fold as the enzyme scaffold may also contribute to the identification of insertion sites for SDRs. For example, haptoglobin (Arcoleo, J. & Greer, J.; (1982) J. *Biol. Chem.* 257, 10063-10068) and azurocidin (Almeida, R. et al. (1991) Biochem. Biophys. Res. Commun. 177, 688-695) share the same chymotrypsin-like fold with all S1 proteases. Due to substitutions in the active site residues these proteins do not posses any proteolytic function, yet they show high homology with active proteases.

Differences between these proteins and specific proteases include regions that can serve as insertion sites for SDRs.

In a fourth approach, insertion sites for SDRs are identified experimentally by techniques such as alanine scanning, random mutagenesis, random insertion or random deletion. In contrast to the approach disclosed above, this approach does not require detailed knowledge about the three-dimensional structure of the scaffold protein. In one preferred variant of this approach, random mutagenesis of enzymes with relatively high specificity from the same structural class as the protein scaffold and screening for loss or change of specificity can be used to identify insertion sites for SDRs in the protein scaffold.

Random mutagenesis, alanine scanning, random insertion or random deletion are all done on the level of the polynucleotides encoding the enzymes. There are a variety of protocols known in the literature (e.g. Sambrook, J.F; Fritsch, E.F.; Maniatis,T.; Cold Spring Harbor Laboratory Press, Second Edition, 1989, New York). For example, random mutagenesis can be achieved by the use of a polymerase as described in patent WO 9218645. According to this patent, the one or more genes encoding the one or more proteases are amplified by use of a DNA polymerase with a high error rate or under conditions that increase the rate of misincorporations. For example the method of Cadwell and Joyce can be employed (Cadwell, R.C. and Joyce, G.F., PCR methods. Appl. 2 (1992) 28-33). Other methods of random mutagenesis such as, but not limited to, the use of mutator stains, chemical mutagens or UV-radiation can be employed as well.

Alternatively, oligonucleotides can be used for mutagenesis that substitute randomly distributed amino acid residues with an alanine. This method is generally referred to as alanine scanning mutagenesis (Fersht, A.R. Biochemistry (1989) 8031-8036). As a further alternative, modifications of the alanine scanning mutagenesis such as binominal mutagenesis (Gregoret, L.M. and Sauer, R.T. PNAS (1993) 4246-4250) or combinatorial alanine scanning (Weiss et al., PNAS (2000) 8950-8954) can be employed.

In order to express engineered enzymes, the DNA encoding such engineered proteins is ligated into a suitable expression vector by standard molecular cloning techniques (e.g. Sambrook, J.F; Fritsch, E.F.; Maniatis, T.; Cold Spring Harbor Laboratory Press, Second Edition, 1989, New York). The vector is introduced in a

suitable expression host cell, which expresses the corresponding engineered enzyme variant. Particularly suitable expression hosts are bacterial expression hosts such as Escherichia coli or Bacillus subtilis, or yeast expression hosts such as Saccharomyces cerevisae or Pichia pastoris, or mammalian expression hosts such as Chinese Hamster Ovary (CHO) or Baby Hamster Kidney (BHK) cell lines, or viral expression systems such as bacteriophages like M13 or Lambda, or viruses such as the Baculovirus expression system. As a further alternative, systems for *in vitro* protein expression can be used. Typically, the DNA is ligated into an expression vector behind a suitable signal sequence that leads to secretion of the enzyme variants into the extracellular space, thereby allowing direct detection of protease activity in the cell supernatant. Particularly suitable signal sequences for Escherichia coli are HlyA, for Bacillus subtilis AprE, NprB, Mpr, AmyA, AmyE, Blac, SacB, and for S. cerevisiae Bar1, Suc2, Matα, Inu1A, Ggplp. Alternatively, the enzyme variants are expressed intracellularly and the substrates are expressed also intracellularly. Preferably, this is done essentially as described in patent application WO 0212543, using a fusion peptide substrate comprising two auto-fluorescent proteins linked by the substrate amino-acid sequence. As a further alternative, after intracellular expression of the enzyme variants, or secretion into the periplasmatic space using signal sequences such as DsbA, PhoA, PelB, OmpA, OmpT or gIII for Escherichia coli, a permeabilisation or lysis step releases the enzyme variants into the supernatant. The destruction of the membrane barrier can be forced by the use of mechanical means such as ultrasonic, French press, or the use of membrane-digesting enzymes such as lysozyme. As another, further alternative, the genes encoding the enzyme variants are expressed cell-free by the use of a suitable cell-free expression system. For example, the S30 extract from Escherichia coli cells is used for this purpose as described by Lesly et al. (Methods in Molecular Biology 37 (1995) 265-278).

The ensemble of gene variants generated and expressed by any of the above methods are analyzed with respect to their affinity, substrate specificity or activity by appropriate assay and screening methods as described in detail for example in patent application PCT/EP03/04864. Genes from catalytically active variants having reduced specificity in comparison to the original enzyme are analyzed by sequencing. Sites at which mutations and/or insertions and/or

deletions occurred are preferred insertion sites at which SDRs can be inserted site-specifically.

In a second embodiment, the one or more fully or partially random peptide sequences are inserted at random sites in the protein scaffold. This modification is usually done on the polynucleotide level, i.e. by inserting nucleotide sequences into the gene that encodes the protein scaffold. Several methods are available that enable the random insertion of nucleotide sequences. Systems that can be used for random insertion are for example ligation based systems (Murakami et al. Nature Biotechnology 20 (2002) 76-81), systems based on DNA polymerisation and transposon based systems (e.g. GPS-M™ mutagenesis system, NEB Biolabs; MGS™ mutation generation system, Finnzymes). The transposon-based methods employ a transposase-mediated insertion of a selectable marker gene that contains at its termini recognition sequences for the transposase as well as two sites for a rare cutting restriction endonuclease. Using the latter endonuclease one usually releases the selection marker and after religation obtains an insertion. Instead of performing the religation one can alternatively insert a fragment that has terminal recognition sequences for one or two outside cutting restriction endonuclease as well as a selectable marker. After ligation, one releases this fragment using the one or two outside cutting endonucleases. After creating blunt ends by standard methods one inserts blunt ended random fragments at random positions into the gene.

In a further preferred embodiment, methods for homologous in-vitro recombination are used to combine the mutations introduced by the above mentioned methods to generate enzyme populations. Examples of methods that can be applied are the Recombination Chain Reaction (RCR) according to patent application WO 0134835, the DNA-Shuffling method according to the patent application WO 9522625, the Staggered Extension method according to patent WO 9842728, or the Random Priming recombination according to patent application WO9842728. Furthermore, also methods for non-homologous recombination such as the Itchy method can be applied (Ostermeier, M. et al. Nature Biotechnology 17 (1999) 1205-1209).

Upon random insertion of a nucleotide sequence into the protein scaffold one obtains a library of different genes encoding enzyme variants. The polynucleotide library is subsequently transferred to an appropriate expression vector. Upon

expression in a suitable host or by use of an in vitro expression system, a library of enzymes containing randomly inserted stretches of amino acids is obtained.

According to step (b) of this third aspect of the invention, one or more fully or partially random peptide sequences are inserted into the protein scaffold. The actual number of such inserted SDRs is determined by the intended quantitative specificity following the relation: the higher the intended specificity is, the more SDRs are inserted. Whereas a single SDR enables the generation of moderately specific enzymes, two SDRs enable already the generation of significantly specific enzymes. However, up to six and more SDRs can be inserted into a protein scaffold. A similar relation is valid for the length of the SDRs: the higher the intended specificity is, the longer are the SDRs that are to be inserted. SDRs can be as short as one to four amino acid residues. They can, however, also be as long as 50 amino acid residues. Significant specificity can already be generated by the use of SDRs of a length of four to six amino acid residues.

The peptid sequences that are inserted can be fully or partially random. In this context, fully random means that a set of sequences are inserted in parallel that includes sequences that differ from each other in each and every position. Partially random means that a set of sequences are inserted in parallel that includes sequences that differ from each other in at least one position. This difference can be either pair-wise or with respect to a single sequence. For example, when regarding an insertion of the length of four amino acids, partial random could be a set (i) that includes AGGG, GVGG, GGLG, GGGI, or (ii) that includes AGGG, VGGG, LGGG and IGGG. Alternatively, random sequences also comprises sequences that differ from each other in length. Randomization of the peptide sequences is achieved by randomization of the nucleotide sequences that are inserted into the gene at the respective sites. Thereby, randomization can be achieved by employing mixtures of nucleobases as monomers during chemical synthesis of the oligonucleotides. A particularly preferred mixture of monomers for a fully random codon that in addition minimizes the probability of stop codons is NN(GTC). Alternatively, random oligonucleotides can be obtained by fragmentation of DNA into short fragments that are inserted into the gene at the respective sites. The source of the DNA to be fragmented may be a synthetic oligonucleotide but alternatively may originate from cloned genes, cDNAs, or

genomic DNA. Preferably, the DNA is a gene encoding an enzyme. The fragmentation can, for example, be achieved by random endonucleolytic digestion of DNA. Preferably, an unspecific endonuclease such as DNAse I (e.g. from bovine pancreas) is employed for the endonucleolytic digestion.

If steps (a) – (c) of the inventive method are repeated cyclically, there are different alternatives for obtaining random peptide sequences that are inserted in consecutive rounds. Preferably, SDRs that were identified in one round as leading to increased specificity of enzyme are used as templates for the random peptide sequences that are inserted in the following round.

In a preferred alternative, the sequences selected in one round are analysed and randomized oligonucleotides are generated based on these sequences. This can, for example, be achieved by using in addition to the original nucleotide with a certain percentage mixtures of the other three nucleotides monomers at each position in the oligonucleotide synthesis. If, for example, in a first round an SDRs is identified that has the amino acid sequence ARLT, e.g. encoded by the nucleotide sequence GCG CGC CTT ACC, a random peptide sequence inserted in this SDR site could be encoded by an oligonucleotide with 70% G, 10% A, 10% T and 10% C at the first position, 70% C, 10% G, 10% T and 10% A at the second position, etc. This leads at each position approximately in 1 of 3 cases to the template amino acid and in 2 of 3 cases to another amino acid.
In another preferred alternative, the sequences selected in one round are analyzed and a consensus library is generated based on these sequences. This can, for example, be achieved by using defined mixtures of nucleotides at each position in the oligonucleotide synthesis in a way that leads to mixtures of the amino acid residues that were identified at each position of the SDR selected in the previous round. If, for example, in a first round two SDRs are identified that have the amino acid sequences ARLT and VPGS, a consensus library inserted in this SDR site in the following round could be encoded by an oligonucleotide with the sequence G(C/T)G C(G/C)C (G/T)(G/T)G (A/T)CC. This would correspond to the random peptide sequence (A/V)(R/P)(L/G/V/W)(T/S), thereby allowing all combinations of the amino acid residues identified in the first round, and, due to the degeneracy of the genetic code, allowing in addition to a lower degree alternative amino acid residues at some positions.

In another preferred alternative, the sequences selected in one round are, without previous analysis, recombined using methods for the in vitro recombination of polynucleotides, such as the methods described in WO 01/34835 (the following also provides details of the eighth and ninth aspect of the invention).

After insertion of the partially or fully random sequences into the gene encoding the scaffold protein, and eventually ligation of the resulting gene into a suitable expression vector using standard molecular cloning techniques (Sambrook, J.F; Fritsch, E.F.; Maniatis,T.; Cold Spring Harbor Laboratory Press, Second Edition, 1989, New York), the vector is introduced in a suitable expression host cell which expresses the corresponding enzyme variant. Particularly suitable expression hosts are bacterial expression hosts such as Escherichia coli or Bacillus subtilis, or yeast expression hosts such as Saccharomyces cerevisae or Pichia pastoris, or mammalian expression hosts such as Chinese Hamster Ovary (CHO) or Baby Hamster Kidney (BHK) cell lines, or viral expression systems such as bacteriophages like M13 T7 phage or Lambda, or viruses such as the Baculovirus expression system. As a further alternative, systems for *in vitro* protein expression can be used. Typically, the DNA is ligated into an expression vector behind a suitable signal sequence that leads to secretion of the enzyme variants into the extracellular space, thereby allowing direct detection of enzyme activity in the cell supernatant. Particularly suitable signal sequences for Escherichia coli are ompA, pelB, HlyA, for Bacillus subtilis AprE, NprB, Mpr, AmyA, AmyE, Blac, SacB, and for S. cerevisiae Bar1, Suc2, Matα, Inu1A, Ggplp. Alternatively, the enzyme variants are expressed intracellularly and the substrates are expressed also intracellularly. According to protease variants this is done essentially as described in patent application WO 0212543, using a fusion peptide substrate comprising two auto-fluorescent proteins linked by the substrate amino-acid sequence. As a further alternative, after intracellular expression of the enzyme variants, or secretion into the periplasmatic space using signal sequences such as DsbA, PhoA, PelB, OmpA, OmpT or gIII for Escherichia coli, a permeabilisation or lysis step releases the enzyme variants into the supernatant. The destruction of the membrane barrier can be forced by the use of mechanical means such as ultrasonic, French press, or the use of membrane-digesting enzymes such as

lysozyme. As another, further alternative, the genes encoding the enzyme variants are expressed cell-free by the use of a suitable cell-free expression system. For example, the S30 extract from Escherichia coli cells is used for this purpose as described by Lesly et al. (Methods in Molecular Biology 37 (1995) 265-278).

After introduction of the vector into host cells, these cells are screened for the expression of enzymes with specificity for the intended target substrate. Such screening is typically done by separating the cells from each other, in order to enable the correlation of genotype and phenotype, and assaying the activity of each cell clone after a growth and expression period. Such separation can for example be done by distribution of the cells into the compartments of sample carriers, e.g. as described in WO 01/24933. Alternatively, the cells are separated by streaking on agar plates, by enclosing in a polymer such as agarose, by filling into capillaries, or by similar methods.
Identification of variants with the intended specificity can be done by different approaches. In the case of proteases, preferably assays using peptide substrates essentially as described in PCT/EP03/04864 are employed.

Regardless of the expression format, selection of enzyme variants is done under conditions that allow identification of enzymes that recognize and convert the target sequence preferably. As a first alternative, enzymes that recognize and convert the target sequence preferably are identified by screening for enzymes with a high affinity for the target substrate sequence. High affinity corresponds to a low K_M which is selected by screening at target substrate concentrations substantially below the K_M of the first enzyme. Preferably, the substrates that are used are linked to one or more fluorophores that enable the detection of the modification of the substrate at concentrations below 10 µM, preferably below 1 µM, more preferably below 100 nM, and most preferably below 10 nM.

As a second alternative, enzymes that recognize and convert the target substrate preferably are identified by employing two or more substrates in the assay and screening for activity on these two or more substrates in comparison. Preferably, the two or more substrates employed are linked to different marker molecules,

thereby enabling the detection of the modification of the two or more substrates consecutively or in parallel. In the case of proteases, particularly preferably two peptide substrates are employed, one peptide substrate having an arbitrarily chosen or even partially or fully random amino-acid sequence thereby enabling to monitor the activity on an arbitrary substrate, and the other peptide substrate having an amino-acid sequence identical to or resembling the intended target substrate sequence thereby enabling to monitor the activity on the target substrate. Especially preferably, these two peptide substrates are linked to fluorescent marker molecules, and the fluorescent properties of the two peptide substrates are sufficiently different in order to distinguish both activities when measured consecutively or in parallel. For example, a fusion protein comprising a first autofluorescent protein, a peptide, and a second autofluorescent protein according to patent application WO 0212543 can be used for this purpose. Alternatively, fluorophores such as rhodamines are linked chemically to the peptide substrates.

As a third alternative, enzymes that recognize and convert the target substrate preferably are identified by employing one or more substrates resembling the target substrate together with competing substrates in high excess. Screening with respect to activity on the substrates resembling the target substrate is then done in the presence of the competing substrates. Enzymes having a specificity which corresponds qualitatively to the target specificity, but having only a low quantitative specificity are identified as negative samples in such a screen. Whereas enzymes having a specificity which corresponds qualitatively and quantitatively to the target specificity are identified positively. Preferably, the one or more substrates resembling the target substrate are linked to marker molecules, thereby enabling the detection of their modifications, whereas the competing substrates do not carry marker molecules. The competing substrates have arbitrarily chosen or random amino-acid sequences, thereby acting as competitive inhibitors for the hydrolysis of the marker-carrying substrates. For example, protein hydrolysates such as Trypton can serve as competing substrates for engineered proteolytic enzymes according to the invention.
As a fourth alternative, enzymes that recognize and convert the target substrate preferably are identified and selected by an amplification-coupled or growth-coupled selection step. Furthermore, the activity can be measured intracellularily

and the selection can be done by a cell sorter, such as a fluorescence-activated cell sorter.

As a further alternative, enzymes that recognize and convert the target substrate are identified by first selecting enzymes that preferentially bind to the target substrate, and secondly selecting out of this subgroup of enzyme variants those enzymes that convert the target substrate. Selection for enzymes that preferentially bind the target substrate can be either done by selection of binders to the target substrate or by counter-selection of enzymes that bind to other substrates. Methods for the selection of binders or for the counter-selection of non-binders is known in the art. Such methods typically require phenotype-genotype coupling which can be solved by using surface display expression methods. Such methods include, for example, phage or viral display, cell surface display and in vitro display. Phage or viral display typically involves fusion of the protein of interest to a viral/phage protein. Cell surface display, i.e. either bacterial or eukaryotic cell display, typically involves fusion of the protein of interest to a peptide or protein that is located at the cell surface. In in-vitro display, the protein is typically made in vitro and linked directly or indirectly to the mRNA encoding the protein (DE 19646372).

The invention also provides for a composition or pharmaceutical composition comprising one or more engineered enzymes according to the first aspect of the invention as defined herein before. The composition may optionally comprise an acceptable carrier, excipient and/or auxiliary agent.

Pharmaceutical compositions according to the invention may optionally comprise a pharmaceutically acceptable carrier. Pharmaceutical formulations are well known and pharmaceutical compositions may be routinely formulated by one having ordinary skill in the art. The composition can be formulated as a solution, suspension, emulsion, or lyophilized powderin association with a pharmaceutically acceptable vehicle. Examples of such vehicles are water, saline, Ringer′s solution, dextrose solution, and human serum albumin. Liposomes and nonaaqueous vehicles such as fixed oils may also be used. The vehicle or lyophilized powder may contain additives that maintain isotonicity (e.g. sodium

chloride, mannitol) and chemical stability (e.g. buffers and preservatives). The composition is sterilized by commonly used techniques.

The pharmaceutical composition of the present invention may be administrated by any means that enables the active agent to reach the agent´s site of action in the body of a mammal. Pharmaceutical compositions may be administered parentally, i.e. intravenous, subcutaneous, intramuscular.
Dosage varies depending upon known factors such as the pharmacodynamic characteristics of the particular agent, and its mode and route of administration ; age, health, and weight of the recipient; nature and extent of symptoms, kind of concurrent treatment, frequency of treatment, and the effect desired.

Non-pharamceutical compositions as defined herein are research composition, nutritional composition, cleaning composition, desinfection composition, cosmetic composition or composition for personal care. Moreover, DNA sequences coding for the engineered enzyme as defined herein before and vectors containing said DNA sequences are also provided. Finally, transformed host cells (prokaryotic or eukaryotic) or transgenic organisms containing such DNA sequences and/or vectors, as well as a method utilizing such host cells or transgenic animals for producing the engineered enzyme of the first aspect of the invention are also contemplated.

Detailed description of the figures
Figure 1: Three-dimensional structure of human trypsin I with the active site residues shown in "ball-and-stick" representation and with the marked regions indicating potential SDR insertion sites.

Figure 2: Alignment of the primary amino acid sequences of the human proteases trypsin I, alpha-thrombin and enteropeptidase all of which belong to the structural class S1 of the serine protease family. Trypsin represents an unspecific protease of this structural class, while alpha-thrombin and enteropeptidase are proteases with high substrate specificity. Compared to trypsin several regions of insertions of three or more amino acids into the primary sequence of a-thrombin and enterokinase are seen. The region marked with (-1-) and the region marked with (-3-) are preferred SDR insertion sites. In

the tertiary structure of alpha-thrombin both regions are in the vicinity of the substrate binding site. These regions therefore fullfil two criteria to be selected as candidates for SDRs: firstly, they represent insertions in the specific proteases compared to the unspecific one and, secondly, they are close to the substrate binding site. A representation of the three-dimensional structure is given in figure 3.

Figure 3: Three-dimensional structure of subtilisin with the active site residues being shown in "ball-and-stick" representation and with the numbered regions indicating potential SDR insertion sites.

Figure 4: Alignment of the primary amino acid sequences of subtilisin E, furin, PC1 and PC5 all of which belong to the structural class S8 of the serine protease family. Subtilisin E represents an unspecific protease of this structural class, while furin, PC1 and PC5 are proteases with high substrate specificity. Compared to subtilisin several regions of insertions of three or more amino acids into the primary sequence of furin, PC1 and PC5 are seen. The regions marked with (-4-), (-5-), (-7-), (-9-) and (-11-) are preferred SDR insertion sites. These regions stretches fulfill two criteria to be selected as candidates for SDRs: firstly, they represent insertions in the specific proteases compared to the unspecific one and, secondly, they are close to the active site residues.

Figure 5: Three-dimensional structure of beta-secretase with the active site residues being shown in "ball-and-stick" representation and with the numbered regions indicating potential SDR insertion sites.

Figure 6: Alignment of the primary amino acid sequences of pepsin, b-secretase and cathepsin D, all of which belong to the structural class A1 of the aspartic protease family. Pepsin represents an unspecific protease of this structural class, while b-secretase and cathepsin D are proteases with high substrate specificity. Compared to pepsin several regions of insertions of three or more amino acids into the primary sequence of b-secretase and cathepsin D are seen. The regions marked with -1- to -11- correspond to possible SDR combining sites and are also marked in Fig.5.

Figure 7: illustrates the three-dimensional structure of caspase 7 with the active site residues being shown in "ball-and-stick" representation and with the numbered regions indicating potential SDR insertion sites.

Figure 8: shows the primary amino acid sequence of caspase 7 as a member of the cysteine protease class C14 family (see also SEQ ID NO: 14).

Figure 9: Schematic representation of method according to the third aspect of the invention.

Figure 10: Western blot analysis of trypsin expression. Supernatant of cell cultures expressing variants of trypsin are compared to negative controls. Lane 1: molecular weight standard; lane 2: negative control; lane 3: supernatant of variant a; lane 4: negative control; lane 5: supernatant of variant b. A primary antibody specific to the expressed protein and a secondary antibody for generation of the signal were used.

Figure 11: Time course of the proteolytic cleavage of a target substrate. Supernatant of cells containing the vector with the gene for human trypsin and that of cells containing the vector without the gene was incubated with the peptide substrate described in the text. Cleavage of the peptide results in a decreased read out value. Proteolytic activity is confirmed for the positive clone.

Figure 12: Relative activity of three engineered proteolytic enzymes in comparison with human trypsin I on two different peptide substrates. A time course of the proteolytic digestion of the two substrates was performed and evaluated. Substrate B was used for screening and substrate A is a closely related sequence. Relative activity of the three variants was normalized to the activity of human trypsin I. Variant 1 and 2 clearly show increased specificity towards the target substrate. Variant 3, on the other hand, serves as a negative control with similar activities as the human trypsin I.

Figure 13: Relative specificities of trypsin and variants of engineered proteolytic enzymes with one or two SDRs, respectively. Activity of the proteases was determined in the presence and absence of competitor substrate, i.e. peptone at a concentration of 10mg/ml. Time courses for the proteolytic cleavage were

recorded and the time constants k determined. The ratios between the time constants with and without competitor were formed and represent a quantitative measure for the specificity of the protease. The ratios were normalized to trypsin. The specificity of the variant containing two SDRs is 2.5 fold higher than that of the variant with SDR2 alone.

Figure 14: Shows the relative specificities of protease variants in absence and presence of competitor substrate. The protease variants containig two inserts with different sequences and the non-modified scaffold human trypsin I were expressed in a suitable host. Activity of the protease variants was determined as the cleavage rate of a peptide with the desired target sequence of TNF-alpha in the absence and presence of competitor substrate. Specificity is expressed as the ratio of cleavage rates in the presence and absence of competitor.

Figure 15: The figure shows the reduction of cytotoxicity induced by human TNF-alpha when incubating the human TNF-alpha with concentrated supernatant from cultures expressing the inventive engineered proteolytic enzymes being specific for human TNF-alpha. This indicates the efficacy of the inventive engineered proteolytic enzymes.

Figure 16: The figure shows the reduction of cytotoxicity induced by human TNF-alpha when incubating the human TNF-alpha with different concentrations of purified inventive engineered proteolytic enzyme being specific for human TNF-alpha. Variant g comprises SEQ ID NO:72 as SDR1 and SEQ ID NO:73 as SDR2. This indicates the efficacy of the inventive engineered proteolytic enzymes.

Figure 17: The figure compares the activity of inventive engineered proteolytic enzymes being specific for human TNF-alpha with the activity of human trypsin I on two protein substrates: (a) human TNF-alpha; (b) mixture of human serum proteins. This indicates the safety of the inventive engineered proteolytic enzymes. Variant x corresponds to Seq ID No: 75 comprising the SDRs according to Seq ID No. 89 (SDR1) and 95 (SDR2). Variants xi and xii correspond to derivatives thereof comprising the same SDR sequences.

Figure 18: Specific hydrolysis of human VEGF by an engineered proteolytic enzyme derived from human trypsin.

Examples

In the following examples, materials and methods of the present invention are provided including the determination of catalytic properties of enzymes obtained by the method. It should be understood that these examples are for illustrative purpose only and are not to be construed as limiting this invention in any manner. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

In the experimental examples described below, standard techniques of recombinant DNA technology were used that were described in various publications, e.g. Sambrook et al. (1989), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, or Ausubel et al. (1987), Current Protocols in Molecular Biology 1987-1988, Wiley Interscience. Unless otherwise indicated, restriction enzymes, polymerases and other enzymes as well as DNA purification kits were used according to the manufacturers specifications.

Example I: Identification of SDR sites in human trypsin

Insertion sites for SDRs have been identified in the serine protease human trypsin I (structural class S1) by comparison with members of the same structural class having a higher sequence specificity. Trypsin represents a member with low substrate specificity, as it requires only an arginine or lysine residue at the P_1 position. On the other hand, thrombin, tissue-type plasminogen activator or enterokinase all have a high specificity towards their substrate sequences, i.e. (L/I/V/F)XPR^ NA, CPGR^ VVGG and DDDK^ , respectively. The primary sequences and tertiary structures of these and further S1 serine proteases have been aligned in order to determine regions of low and high sequence and structure homology and especially regions that correspond to insertions in the sequences of the more specific proteases (Figure 2). Several regions of insertions equal or longer than 3 amino acids representing potential SDR sites have been identified as indicated in Figure 1. These regions were chosen as target sites for the insertion of SDRs in the examples below, e.g. SDR1 (region one in figure 2, after amino acid 42 according to SEQ ID NO:1) with a length of six and SDR2 (region three in figure 2, after amino acid 123 according to SEQ ID NO:1) with a length of five amino acids, respectively.

Example II: Molecular cloning of the human trypsin I gene to be used as scaffold protein and expression of the mature protease in B. subtilis

The gene encoding the unspecific protease human trypsinogen I was cloned into the vector pUC18. Cloning was done as follows: the coding sequence of the protein was amplified by PCR using primers that introduced a KpnI site at the 5' end and a BamHI site at the 3' end. This PCR fragment was cloned into the appropriate sites of the vector pUC18. Identity was confirmed by sequencing. After sequencing the coding sequence of the mature protein was amplified by PCR using primers that introduced different BglI sites at the 5' end and the 3' end.

This PCR fragment was cloned into the appropriate sites of an E. coli – B. subtilis shuttle vector. The vector contains a pMB1 origin for amplification in E. coli, a neomycin resistance marker for selection in E. coli, as well as a P43 promoter for the constitutive expression in B. subtilis. A 87 bp fragment that contains the leader sequence encoding the signal peptide from the sacB gene of B. subtilis was introduced behind the P43 promoter. Different BglI restriction sites serve as insertion sites for heterologous genes to be expressed.

Expression of human trypsin I was confirmed by measurement of the proteolytic aciticity in supernatant of cells containing the vector with the gene in comparison to a negative control. A peptide including an arginine cleavage site was chosen as a substrate. The peptide was N-terminally biotinylated and labeled with a fluorophore at the C-terminus. After incubation of the peptide with culture supernatant streptavidin was added. Uncleaved peptide associate with streptavidin and lead to a high read out value while cleavage results in low read out values. Figure 11 shows the time course of a proteolytic digestion of B. subtilis cells containing the vector with the trypsin I gene in comparison to B. subtilis cells containing the vector without the trypsin I gene (negative control).

As a further confirmation of expression of the protease, supernatants of cells containing the vector with the gene and control cells were analyzed by polyacrylamid gel electrophoreses and subsequent western blot using an antibody specific to the target protease. The procedure was performed according to standard methods (Sambrook, J.F; Fritsch, E.F.; Maniatis,T.; Cold Spring Harbor Laboratory Press, Second Edition, 1989, New York). Figure 8 confirms expression of the protein only in the cells harbouring the vector with the gene for trypsin.

Example III: Providing a scaffold protein

In this example, human trypsin I was used as the scaffold protein. The gene was either used in its natural form, or, alternatively, was modified to result in a scaffold protein with increased catalytic activity or further improved characteristics.

The modification was done by random modification of the gene, followed by expression of the enzyme and subsequent selection for increased activity. First, the gene was PCR amplified under error-prone conditions, essentially as described by Cadwell, R.C and Joyce, G.F. (PCR Methods Appl. 2 (1992) 28-33). Error-prone PCR was done using 30 pmol of each primer, 20 nmol dGTP and dATP, 100 nmol dCTP and dTTP, 20 fmol template, and 5 U Taq DNA polymerase in 10 mM Tris HCl pH 7.6, 50 mM KCl, 7 mM MgCl2, 0.5 mM MnCl2, 0.01 % gelatin for 20 cycles of 1 min at 94 °C, 1 min at 65 °C and 1 min at 72 °C. The resulting DNA library was purified using the Qiaquick PCR Purification Kit following the suppliers' instructions. The PCR product was digested with the restriction enzyme *BglI* and purified. Afterwards, the PCR product was ligated into the E. coli – B. subtilis shuttle vector described above which was digested with BglI and dephosphorylated. The ligation products were transformed into E. coli, amplified in LB, and the plasmids were purified using the Qiagen Plasmid Purification Kit following the suppliers' instructions. Resulting plasmids were transformed into B. subtilis cells.

Alternatively, or in addition to random mutagenesis, variants of the gene were statistically recombined at homologous positions by use of the Recombination Chain Reaction, essentially as described in WO 0134835. PCR products of the genes encoding the protease variants were purified using the QIAquick PCR Purification Kit following the suppliers' instructions, checked for correct size by agarose gel electrophoresis and mixed together in equimolar amounts. 80 µg of this PCR mix in 150 mM TrisHCl pH 7.6, 6.6 mM $MgCl_2$ were heated for 5 min at 94 °C and subsequently cooled down to 37 °C at 0.05 °C/s in order to re-anneal strands and thereby produce heteroduplices in a stochastic manner. Then, 2.5 U Exonuclease III per µg DNA were added and incubated for 20, 40 or 60 min at 37 °C in order to digest different lengths from both 3' ends of the heteroduplices. The partly digested PCR products were refilled with 0.6 U Pfu polymerase per µg DNA by incubating for 15 min at 72 °C in 0.17 mM dNTPs and Pfu polymerase

buffer according to the suppliers' instructions. After performing a single PCR cycle, the resulting DNA was purified using the QIAquick PCR Purification Kit following the suppliers' instructions, digested with BglI and ligated into the linearized vector. The ligation products were transformed into E. coli, amplified in LB containing ampicillin as marker, and the plasmids were purified using the Qiagen Plasmid Purification Kit following the suppliers' instructions. Resulting plasmids were transformed into B. subtilis cells.

Example IV: Insertion of SDRs into the protein scaffold of human trypsin I and generation of an engineered proteolytic enzyme with specificity for a peptide substrate having the sequence KKWLGRVPGGPV.

In order to create insertion sites for SDRs in human trypsin I, two pairs of different restriction sites were introduced into the gene at sites that were identified as potential SDR sites (see Example I above) without changing the amino acid sequence. The insertion of the restriction sites was done by overlap extension PCR. Primers restr1 and restr2 were used for the introduction of SacII and BamHI restriction sites, restr3 and restr4 were used for the introduction of KpnI and NheI restriction sites. The sequences of the primers were as follows:

Binding site for restr1 and restr2 and the corresponding amino acid sequence (SEQ ID NO:54):

5'-GGTGGTATCAGCAGGCCACTGCTACAAGTCCCGCATCCAGGT-3'
V V S A G H C Y K S R I Q

Forward primer restr1 (SEQ ID NO:56):

5'-GGTGGTATCCGCGGGCCACTGCTACAAGTCCCGGATCCAGGT-3'

Reverse primer restr2 (SEQ ID NO:57):

5'-ACCTGGATCCGGGACTTGTAGCAGTGGCCCGCGGATACCACC-3'

Binding site for restr3 and restr4 and the corresponding amino acid sequence (SEQ ID NO:58):

5'-CCACTGGCACGAAGTGCCTCATCTCTGGCTGGGGCAACACTGCGAGCTCT-3'
T G T K C L I S G W G N T A S S

Forward primer restr3 (SEQ ID NO:60):

5'-CCACTGGCACGAAGTGCCTCATCTCTGGCTGGGGCAACACTGCGAGCTCT-3'

Reverse primer restr4 (SEQ ID NO:61):

5'-AGAGCTAGCAGTGTTGCCCCAGCCAGAGATGAGGCACTTGGTACCAGTGG-3'

In a first overlap extension PCR, the SacII/BamHI sites were introduced, enabling to insert SDR1, and in a second overlap extension PCR the KpnI/NheI sites, enabling the insertion of SDR2. The product of the overlap extension PCR was amplified using primers pUC-forward and pUC-reverse. The sequences of pUC-forward and pUC-reverse are as follows:

pUC-forward (SEQ ID NO:62): 5'-GGGGTACCCCACCACCATGAATCCACTCCT-3'

pUC-reverse (SEQ ID NO:63): 5'-CGGGATCCGGTATAGAGACTGAAGAGATAC-3'

The restriction sites generated thereby were subsequently used to insert defined or random oligonucleotides into the SDR1 and SDR2 insertion sites by standard restriction and ligation methods. Typically, two complementary synthetic 5'-phosphorylated oligonucleotides were annealed and ligated into a vector carrying the modified human trypsin I gene that was cleaved with the respective restriction enzymes. Oligonucleotides encoding SDR1 were inserted via the SacII/BamHI sites whereas oligonucleotides encoding SDR2 were inserted via the KpnI/NheI sites. For each insertion an oligonucleotide pair according to the following general sequences was used ([P] indicating 5'-phosphorylation, N and X indicating any nucleotide or amino acid residue, respectively):

oligox-SDR1f (SEQ ID NO:64):

5'-[P]-GGGCCACTGCTACNNNNNNNNNNNNNNNNNNAAGTCCCG-3'

oligox-SDR1r (SEQ ID NO:66):

3'-CGCCCGGTGACGATGNNNNNNNNNNNNNNNNNNTTCAGGGCCTAG-[P]-5'

G H C Y X X X X X X K S

oligox-SDR2f (SEQ ID NO:67):

5'-[P]-CAAGTGCCTCATCTCTGGCTGGGGCAACNNNNNNNNNNNNNNNACTG-3'

oligox-SDR2r (SEQ ID NO:69):

3'-CATGGTTCACGGAGTAGAGACCGACCCCGTTGNNNNNNNNNNNNNNNNTGACGATC-[P]-5'

K C L I S G W G N X X X X X T

As an alternative to the above method, a PCR based method was used for the integration of random-sequences into the SDR1 and SDR2 insertion sites in the modified human trypsin I. For each SDR, one primer was used where the SDR region is fully randomized. Sequences of the primers were as follows (N = A/C/G/T, B = C/G/T, V = A/C/G):

Primer SDR1-mutnnb-forward (SEQ ID NO:70):

5'-TGGTATCCGCGGGCCACTGCTACNNBNNBNNBNNBNNBNNBAAGTCCCGGATCCAGGTG-3'

Primer SDR2-mutnnb-reverse (SEQ ID NO:71):

5'-GGCGCCAGAGCTAGCAGTVNNVNNVNNVNNVNNGTTGCCCCAGCCAGAGATG-3'

The codon NNB, or VNN in the reverse strand, allows all 20 amino acids to made, but reduces the probability of encoding a stop codon from 0.047 to 0.021.

As a further alternative, after identification of SDRs that lead to increased specificity, these SDRs were used as templates for further randomization. Thereby, random peptide sequences were inserted that were partially randomized at each position and partially identical at each position to the original sequence.

As an example, random peptide sequences that have in approximately 1 of 3 cases the template amino acid residue and in approximately 2 of 3 cases any other amino acid residue at each position were inserted into the two SDR insertion sites of the modified human trypsin I. For this purpose, primers that contain at each nucleotide position of the SDR approximately 70% of the template bases and 30% of a mixture of the three other bases were used.

With each primer pair a PCR was performed under standard conditions using the human trypsin I gene as template. The resulting DNA was purified using the QIAquick PCR Purification Kit following the suppliers' instructions and digested with SacII and NheI. After digestion the DNA was purified and ligated into the SacII and NheI digested and dephosphorylayted vector. The ligation products were transformed into E. coli, amplified in LB containing the respective marker, and the plasmids were purified using the Qiagen Plasmid Purification Kit following the suppliers' instructions. Resulting plasmids were transformed into B. subtilis cells. These cells were then separated to single cells, grown to clones, and after expression of the protease gene screened for proteolytic activity.

The following substrates were employed for screening for proteolytic activity (SEQ ID NOs:76 and 77):

substrate A	L	L	W	L	G	R	V	V	G	G	P	V
substrate B	K	K	W	L	G	R	V	P	G	G	P	V

Protease variants were screened on substrate B at complexities of 10^6 variants by confocal fluorescence spectroscopy. The substrate was a peptide biotinylated at the N-terminus and fluorescently labeled at the C-terminus. After incubation of the peptide with supernatant of cells expressing different variants of the protease, streptavidin is added and the samples are analysed by confocal fluorimetry. The low concentration of the peptide (20nM) leads to a preferential cleavage by proteases with a high k_{cat}/K_M value, i.e. proteases with high specificity towards the target sequence.

Variants selected in the screening procedure were further evaluated for their specificity towards substrate B and closely related substrate A by measuring time courses of the proteolytic digestion and determining the rate constants which are proportional to the k_{cat}/K_M values. Clearly, compared to the human trypsin that was used as scaffold protein, the specific activity of variants 1 and 2 is shifted (SEQ ID NOs: 2 and 3, respectively) towards substrate B. Variant 3 (SEQ ID NO:4), on the other hand, serves as a negative control with similar activities as the human trypsin I. Sequencing of the genes of the three variants revealed the following amino acid sequences in the SDRs.

Table 2: Sequences of the two SDRs in three different variants selected for specific hydrolysis of substrate B (SEQ ID NOs:78-83).

	SDR 1						SDR 2				
Trypsin	-	-	-	-	-	-	-	-	-	-	-
Variant 1	D	A	V	G	R	D	T	I	T	N	S
Variant 2	N	G	R	D	L	E	V	R	G	T	W
Variant 3	G	F	V	M	F	N	R	S	P	L	T

In a further experiment a pool of variants containing different numbers of SDRs per gene were screened for increased specificity using a mixture of the defined substrate and pepton as a competing substrate. Variants containing one or two

SDRs per gene have been analyzed further. As a measure for the specificity the activity in the peptide cleavage assay was compared with and without the presence of the competing substrate. The concentration of the competing substrate was 10mg/ml. Under these conditions, unspecific proteases show, compared to specific proteases, a stronger decrease in activity with increasing competitor concentrations (range between 0 and 100mg/ml). The ratio of proteolytic activity with and without substrate is a quantitative measure for the specificity of the proteases. Figure 9 shows the relative activities with and without competing substrate. Human trypsin I that was used as the scaffold protein and two variants, one containing only SDR2, and one containing both SDRs, were compared. The specificity of the variant with both SDRs is by a factor of 2.5 higher than that of the variant with SDR2 only, confirming that there is a direct relation between the number of SDRs and the quantitative specificity of resulting engineered proteolytic enzymes.

Example V: Generation of an engineered proteolytic enzyme that specifically inactivates human TNF-alpha

Human trypsin alpha I or a derivative comprising one or more of the following amino acid substitutions E56G; R78W; Y131F; A146T; C183R was used as protein scaffold for the generation of an engineered proteolytic enzyme with high specificity towards human TNF-alpha. The identification of SDR sites in human trypsin I or derivatives thereof was done as described above. Two insertion sites within the scaffold were choosen for SDRs. The protease variants containing two inserts with different sequences and also the human trypsin I itself with no inserts were expressed in a *Bacillus subtilis* cells. The variant protease cells were separated to single cell clones and the protease expressing variants were screened for proteolytic activity on peptides with the desired target sequence of TNF-alpha. The activity of the protease variants was determined as the cleavage rate of a peptide with the desired target sequence of TNF-alpha in the absence and presence of competitor substrate. The specificity is expressed as the ratio of cleavage rates in the presence and absence of competitor (Fig. 14).

Table 3: Relative specificity of variants of engineered proteolytic enzymes with different SDR sequences in absence and presence of competitor substrate (SEQ ID NOs:84-95).

	k with comp./ k without comp	Seq. of SDR 1	Seq. of SDR 2

scaffold (no SDRs)	0.092	---	----
variant a	0.130	RPWDPS	VHPTS
variant b	0.187	GFVMFN	RSPLT
variant c	0.235	EIANRE	RGART
variant d	0.310	KAVVGT	RTPIS
variant e	0.374	VNIMAA	TTARK
variant f	0.487	AAFNGD	RKDFW

The antagonistic effect of three inventive protease variants on human TNF-alpha is shown in Figure 15. By the use of the variants, the induction of apoptosis is almost completely eliminated indicating the anti-inflammatory efficacy of the inventive proteases to initiate TNF-alpha break down. TNF-alpha has been incubated with concentrated supernatant from cultures expressing the variants i to iii for 2 hours. The resulting TNF-alpha has been incubated with non-modified cells for 4 hours. The effect of the remaining TNF-alpha activity was determined as the extent of apoptosis induction by detection of activated caspase-3 as marker for apoptotic cells. For the controls either no protease was added with the human TNF-alpha (dead cells) or buffer instead of human TNF-alpha (live cells) was used, respectively. An analogous experiment is shown in Figure 16 using purified variant xiii. TNF-alpha was incubated with different concentrations of the purified inventive protease variant.

To demonstrate the specificity of the inventive protease variants, proteins from human blood serum or purified human TNF-alpha have been incubated with human trypsin I or the inventive engineered proteolytic enzyme variants, respectively. Here, variant x corresponds to Seq ID No: 75 comprising the same SDRs as variant f, i.e. SDRs according to Seq ID No. 89 (SDR1) and 95 (SDR2). Variants xi and xii correspond to derivatives thereof comprising the same SDR sequences. Remaining intact protein was was determined as a function of time. While the variants as well as human trypsin I digest human TNF-alpha, only trypsin shows activity on serum protein (Figure 17 a and b). This demonstrates the high TNF-alpha specificity of the inventive proteolytic enzymes and indicates their safety and accordingly their low side effects for therapeutic use.

Example VI: Generation of an engineered proteolytic enzyme that specifically hydrolysis human VEGF.

Human trypsin I was used as protein scaffold for the generation of an engineered proteolytic enzyme with high specificity towards human VEGF. The identification of SDR sites in human trypsin I was done as described above. Two insertion sites within the scaffold were choosen for SDRs. The protease variants containing two inserts with different sequences were expressed in *Bacillus subtilis* cells. The variant protease cells were separated to single cell clones and the protease expressing variants were screened as described above. The activity of the protease variants was determined as the rate of VEGF cleavage. 4µg of recombinant human VEGF165 was incubated with 0.18 µg of purified protease in PBS / pH 7.4 at room temperature. Aliquots were taken at the indicated time points and analysed on a polyacrylamide gel. The extend of cleavage was quantified by densitometric analysis of the bands. The activity is plotted over incubation time in Figure 18. Specific cleavage was controlled by further SDS polyacrylamide gel analyses.

Claims

1. Use of a protease with defined specificity for a target substrate for preparing a medicament for the treatment of a specific disease related to said target substrate.

2. The use according to claim 1, wherein
(I) the protease hydrolizes the target substrate and thereby
(i) eliminates or reduces one or more biological activities or physico-chemical properties or pharmacological properties of the target protein, and/or
(ii) activates or increases one or more biological activities or physico-chemical properties or pharmacological properties of the target protein, and/or
(iii) adds one or more biological activities or physico-chemical properties or pharmacolocical properties to the target protein;
and/or
(II) the target substrate hydrolyzed by the protease is a soluble protein, in particular a cytokine, such as the TNF-superfamily proteins, interleukines, interferons, chemokines and growth factors; a hormone; a toxin; an enzyme, such as oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases; a structural protein, such as collagens; or an immunoglobulin; or a membrane associated protein, in particular a single pass transmembrane protein; a multipass transmembrane protein, such as G-protein coupled receptors, ion channels and transporters; a lipid-anchored membrane protein or a GPI-anchored membrane protein.

3. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is a TNF-superfamily protein.

4. The use according to claim 3, wherein
(i) the protease is capable of hydrolysing peptide bonds in human tumor necrosis factor-alpha (hTNF-α, SEQ ID NO:96) or related molecules of the same structure class, preferably the peptide bonds between positions 31/32, 32/33, 44/45, 45/46, 87/88, 128/129, 130/131, 140/141 and/or 141/142, most preferably

between positions 31/32, 32/33 and/or 45/46 of hTNF-α or between analogous positions in said related molecules; and/or
(i) the medicament is suitable for the treatment of rheumatoid arthritis, inflammatory bowel diseases, psoriasis, Crohn's disease, Ulcerative colitis, diabetes type II, classical Hodgkin's Lymphoma (cHL), Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, systemic lupus erythematosus, multiple sclerosis, Systemic inflammatory response syndrome (SIRS), multiple organ dysfunction syndrome (MODS), eosinophilia, neurodegenerative disease, stroke, closed head injury, encephalitis, CNS disorders, asthma, rheumatoid arthritis, sepsis, vasodilation, intravascular coagulation, multiple organ failure or other diseases connected with hTNF-α.

5. The use according to claim 4, wherein the protease is derived from a serine protease of the structural class S1, and preferably wherein the protease is derived from human trypsin I (SEQ ID NO:1).

6. The use according to claim 4, wherein the protease has the a sequence shown in SEQ ID NO:74 or 75 or a derivative thereof and is capable of hydrolysing hTNF-α at positions 31/32 and/or 32/33.

7. The use according to claim 3, wherein
(i) the protease is capable of hydrolysing peptide bonds in human Tumor necrosis factor ligand superfamily member 5 (CD40-L) or related molecules of the same structural class, preferably the peptide bonds between positions 117/118, 133/134, 145/146, 165/166, 200/201, 201/202, 207/208, 216/217 and/or 243/244, most preferred between positions 133/134, 165/166, 201/202 and/or 216/217 of CD40-L (SEQ ID NO:143), or between analogous positions in related molecules; and/or
(ii) the medicament is suitable for the treatment of systemic lupus erythematosus, classical Hodgkin's Lymphoma (cHL) or other diseases connected with CD40-L.

8. The use according to claim 3, wherein
(i) the engineered protease is capable of hydrolysing peptide bonds in human Macrophage migration inhibitory factor (hMIF) or related molecules of the same

structural class, preferably the peptide bonds between positions 16/17, 44/45, 66/67, 73/74, 77/78, 88/89, 92/93 and/or 100/101, most preferred between positions 16/17 and/or 92/93 of hMIF (SEQ ID NO:109) or between analogous positions in related molecules; and/or
(ii) the medicament is suitable for the treatment of inflammatory diseases, or other diseases connected with hMIF.

9. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is an interleukine.

10. The use according to claim 9, wherein
(i) the protease is capable of hydrolysing peptide bonds in Interleukin-1 beta (IL-1ß) or related molecules of the same structural class, preferably the peptide bonds between positions 24/25, 35/36, 46/47, 54/55, 74/75, 75/76, 76/77, 77/78, 86/87, 88/89, 93/94, 94/95, 97/98 and/or 150/151, most preferred between positions 35/36, 75/76, 76/77, 88/89, 93/94, 94/95 and/or 150/151 of IL-1ß (SEQ ID NO:112), or between analogous positions in related molecules; and/or
(iii) the medicament is suitable for the treatment of diabetes, brain inflammation in cancer, arthritis, autoimmune and inflammatory diseases or other diseases connected with IL-1ß.

11. The use according to claim 9, wherein
(i) the protease is capable of hydrolysing peptide bonds in human Interleukin 2 (hIL-2) or related molecules of the same structural class, preferably the peptide bonds between positions 20/21, 32/33, 38/39, 43/44, 45/46 48/49, 49/50, 54/55, 64/65, 76/77, 83/84, 84/85, 107/108, 109/110 and/or 120/121, most preferred between positions 109/110 of hIL-2 (SEQ ID NO:99) or between analogous positions in related molecules; and/or
(ii) the medicament is suitable for the treatment of T-cell leukemia, hairy cell leukemia, Crohn's disease, Ulcerative colitis, Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, systemic lupus erythematosus, multiple sclerosis, asthma and chronic obstructive pulmonary, classical Hodgkin's Lymphoma (cHL) or other diseases connected with hIL-2.

12. The use according to claim 9, wherein
(i) the protease is capable of hydrolysing peptide bonds in human Interleukin 3 (hIL-3) or related molecules of the same structural class, preferably peptide bonds between positions 21/22, 28/29, 36/37, 44/45, 46/47, 51/52, 63/64, 66/67, 79/80, 94/95, 101/102, 108/109 and/or 109/110, most preferred between positions 21/22, 28/29, 46/47, 63/64, 66/67, 79/80 and/or 101/102 of hIL-3 (SEQ ID NO:148), or between analogous positions in related molecules; and/or
(ii) the medicament is suitable for the treatment of classical Hodgkin's Lymphoma (cHL), eosinophilia or other diseases connected with hIL-3.

13. The use according to claim 9, wherein
(i) the protease is capable of hydrolysing peptide bonds in human Interleukin 4 (hIL-4) or related molecules of the same structural class, preferably the peptide bonds between positions 4/5, 12/13, 31/32, 37/38, 61/62, 62/63, 64/65, 91/92, 102/103, 121/122 and/or 126/127, most preferred between positions 4/5, 61/62, 62/63, 64/65 and/or 121/122 of hIL-4 (SEQ ID NO:118), or between analogous positions in related molecules; and/or
(ii) the medicament is suitable for the treatment of classical Hodgkin's Lymphoma (cHL), Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, Asthma, chronic obstructive pulmonary disease, allergic inflammatory reactions or other diseases connected with hIL-4.

14. The use according to claim 9, wherein
(i) the protease is capable of hydrolysing peptide bonds in human Interleukin-5 (hIL-5) or related molecules of the same structural class, preferably the peptide bonds between positions 12/13, 32/33, 67/68, 76/77, 77/78, 80/81, 83/84, 84/85, 85/86, 90/91, 91/92, 92/93 and/or 98/99, most preferred between positions 90/91, 91/92, 92/93 and/or 98/99 of hIL-5 (SEQ ID NO:133), or between analogous positions in related molecules; and/or
(ii) the medicament is suitable for the treatment of classical Hodgkin's Lymphoma (cHL), asthma, chronic obstructive pulmonary disease,

eosinophilia, allergic inflammatory diseases or other diseases connected with hIL-5.

15. The use according to claim 9, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Interleukin-6 (hIL-6) or related molecules of the same structural class, preferably the peptide bonds between positions 32/33, 35/36, 55/56, 71/72, 129/130, 130/131, 132/133, 135/136, 141/142, 161/162, 180/181 and/or 183/184, most preferred between positions 135/136 and/or 141/142 of hIL-6 (SEQ ID NO:134), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of classical Hodgkin's Lymphoma (cHL), breast cancer, renal cell carcinoma, multiple myeloma, lymphoma, leukemia, Grave's disease, Hashimoto's thyroiditis, Sjogren's syndrome, systemic lupus erythematosus, Systemic inflammatory response syndrome (SIRS) which leads to distant organ damage and multpile organ dysfunctionsyndrome (MODS), chronic obstructive pulmonary disease (COPD), Castleman's diseases, inflammatory bowel diseases, Crohn's disease or other diseases connected with hIL-6.

16. The use according to claim 9, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Interleukin 8 (hIL-8) or related molecules of the same structural class, preferably the peptide bonds between positions 11/12, 15/16, 45/46, 47/48, 52/53, 54/55, 60/61, 64/65 and/or 67/68, most preferred between positions 45/46 of hIL-8 (SEQ ID NO:100), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of Crohn's disease, Ulcerative colitis, classical Hodgkin's Lymphoma (cHL), Systemic inflammatory response syndrome (SIRS) which leads to distant organ damage and multpile organ dysfunction syndrome (MODS), chronic obstructive pulmonary disease (COPD), endometriosis, psoriasis, atherosclerotic lesions or other diseases connected with hIL-8.

17. The use according claim 9, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Interleukin-10 (hIL-10) or related molecules of the same structural class, preferably the peptide bonds between positions 24/25, 25/26, 27/28, 28/29, 40/41, 44/45, 49/50, 57/58, 59/60, 84/85, 86/87, 106/107, 107/108, 110/111, 130/131, 134/135, 137/138, 138/139 and/or 144/145, most preferred between positions 24/25, 27/28, 44/45, 49/50, 86/87, 137/138 and/or 144/145 of hIL-10 (SEQ ID NO:135), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of classical Hodgkin's Lymphoma (cHL), diseases related to the suppression of cytotoxic T-cells or other diseases connected with hIL-10.

18. The use according to claim 9, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Interleukin 12 beta chain (hIL-12ß) or related molecules of the same structural class, preferably the peptide bonds between positions 14/15, 18/19, 29/30, 34/35, 87/88, 99/100, 102/103, 104/105, 161/162, 174/175, 222/223, 225/226, 228/229, 238/239, 268/269 and/or 293/294, most preferred between positions 18/19, 34/35, 87/88 and/or 161/162 of hIL-12ß (SEQ ID NO:97) or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of Crohn's disease, classical Hodgkin's Lymphoma (cHL) or other diseases connected with hIL-12ß.

19. The use according to claim 9, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Interleukin 13 (hIL-13) or related molecules of the same structural class, preferably the peptide bonds between positions 25/26, 62/63, 65/66, 86/87, 87/88, 98/99, 108/109 and/or 111/112, most preferred between positions 87/88 of hIL-13 (SEQ ID NO:119), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of cancer, classical Hodgkin's Lymphoma (cHL), eosinophilia, asthma, chronic obstructive pulmonary disease, fibrosis, psoriasis, atopic dermatitis, Ulcerative colitis or other diseases connected with hIL-13.

20. The use according to claim 9, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Interleukin 18 (hIL-18) or related molecules of the same structural class, preferably the peptide bonds between positions 17/18, 32/33, 37/38, 39/40, 40/41, 53/54, 58/59, 79/80, 90/91, 93/94, 98/99, 110/111, 120/121, 123/124, 131/132, 132/133, 142/143, 147/148 and/or 157/158, most preferred between positions 37/38, 132/133, 142/143 and/or 157/158 of hIL-18 (SEQ ID NO:98) or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of Crohn's disease, inflammation liver injuries, pulmonary tuberculosis, plural tuberculosis, rheumatoid arthritis or other diseases connected with hIL-18.

21. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is an interferone.

22. The use according to claim 21, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Interferon-gamma (hIFN-?) or related molecules of the same structural class, preferably the peptide bonds between positions 2/3, 6/7, 13/14, 21/22, 24/25, 34/35, 36/37, 37/38, 62/63, 68/69, 83/84, 86/87, 90/91, 102/103, 107/108 and/or 108/109, most preferred between positions 13/14, 24/25, 37/38, 62/63, 68/69, 102/103 and/or 107/108 of hIFN-? (SEQ ID NO:137), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of classical Hodgkin's Lymphoma (cHL), Crohn's disease, type I diabetes or other diseases connected with IFN-?.

23. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is a chemokine.

24. The use according to claim 23, wherein

(i) the protease is capable of hydrolysing peptide bonds in human small inducible cytokine A2 (hCCL2) or related molecules of the same structural class, preferably the peptide bonds between positions 3/4, 13/14, 18/19, 19/20, 24/25, 29/30, 38/39, 54/55, 56/57, 58/59, 62/63, 65/66 and/or 68/69, most preferred between positions 19/20, 29/30, 38/39, 54/55 and/or 62/63 of hCCL2 (SEQ ID NO:102), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of Crohn's disease, Ulcerative colitis, or other diseases connected with hCCL2.

25. The use according to claim 23, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Eotaxin (hCCL11) or related molecules of the same structural class, preferably the peptide bonds between positions 11/12, 16/17, 17/18, 22/23, 27/28, 33/34, 44/45, 47/48, 48/49, 52/53, 54/55, 56/57, 60/61, 66/67, and/or 73/74, most preferred between positions 48/49 and/or 66/67 of hCCL11 (SEQ ID NO:101), or between analogous positions in related molecules; and/or

(ii) wherein the medicaments is suitable for the treatment of Crohn's disease and Ulcerative colitis, classical Hodgkin's Lymphoma (cHL), chronic pathophysiologic dysfunction, characterized by an influx mainly of Th2 cells, eosinophilia or other diseases connected with hCCL11.

26. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is a growth factor.

27. The use according to claim 26, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Vascular endothelial growth factor (hVEGF) or related molecules of

the same structural class, preferably the peptide bonds between positions 16/17, 19/20, 23/24, 34/35, 41/42, 56/57, 62/63, 63/64, 64/65, 65/66, 82/83, and/or 84/85, most preferred between positions 23/24, 41/42, 63/64, 82/83 and/or 84/85 of hVEGF (SEQ ID NO:103), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of all solid tumors and metastatic solid tumors, inflammatory breast cancer or other diseases connected with hVEGF.

28. The use according to claim 26, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Transforming growth factor beta 1 (hTGF-ß1) or related molecules of the same structural class, preferably the peptide bonds between positions 23/24, 25/26, 26/27, 27/28, 37/38, 55/56 and/or 94/95, most preferred between positions 25/26, 55/56 and/or 94/95 of hTGF-ß1 (SEQ ID NO:104), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of a variety of cancers including breast cancer, colorectal cancer and classical Hodgkin's Lymphoma (cHL); fibrosis, suppression of cell-mediated immunity, glaucoma, diffuse systemic sclerosis or other diseases connected with hTGF-ß1.

29. The use according to claim 26, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Somatotropin (hGrowth hormone; hGH) or related molecules of the same structural class, preferably the peptide bonds between positions 8/9, 16/17, 19/20, 26/27, 33/34, 38/39, 41/42, 70/71, 77/78, 94/95, 103/104, 112/113, 115/116, 116/117, 130/131, 147/148, 154/155 and/or 178/179, most preferred between positions 112/113, 147/148 and/or 154/155 of GH (SEQ ID NO:121), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of acromegaly, diabetes and diabetic kidney disease including renal hypertrophy and glomerular enlargement, cardiovascular disorders or other diseases connected with hGH.

30. The use according to claim 26, wherein

(i) the protease is capable of hydrolysing peptide bonds in Insulin-like growth factor II (hIGF-II) or related molecules of the same structural class, preferably the peptide bonds between positions 15/16, 23/24, 24/25, 34/35, 37/38, 38/39, 48/49 and/or 49/50, most preferred between positions 23/24 of hIGF-II (SEQ ID NO:122), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of diabetes, diabetic kidney disease or other diseases connected with hIGF-II.

31. The use according to claim 26, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Hepatocyte growth factor (hHGF) or related molecules of the same structural class, preferably the peptide bonds between positions 54/55, 60/61, 62/63, 63/64, 68/69, 76/77, 112/113, 123/124, 134/135, 168/169, 198/199 and/or 202/203, most preferred between positions 63/64, 68/69, 76/77, 168/169 and/or 202/203 of hHGF (SEQ ID NO:120), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of angiogenic disorders, hepatocellular carcinoma or other diseases connected with hHGF.

32. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is soluble hormone.

33. The use according to claim 32, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Insulin (hInsulin) or related molecules of the same structural class,

preferably the peptide bonds between positions 16/17 and/or 22/23 of hInsulin B chain (SEQ ID NO:105), and/or between positions 14/15 of Insulin A chain (SEQ ID NO:106) or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of insulin overdosage or other diseases connected with hInsulin.

34. The use according to claim 32, wherein

(i) the engineered protease is capable of hydrolysing peptide bonds in human Ghrelin (Ghrelin) or related molecules of the same structural class, preferably the peptide bonds between positions 1/2, 2/3, 3/4 and/or 4/5 of hGhrelin (SEQ ID NO:107), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of obesity or other diseases connected with hGhrelin.

35. The use according to claim 32, wherein

(i) the protease is capable of hydrolysing peptide bonds in human angiotensinogen (angiotensin) or related molecules of the same structural class, preferably the peptide bonds between positions 1/2, 3/4, and/or 7/8, most preferred between positions 3/4 of angiotensin (SEQ ID NO:108), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of essential hypertension and other diseases connected with angiotensin.

36. The use according to claim 32, wherein

(i) the engineered protease is capable of hydrolysing peptide bonds in human leptin (leptin) or related molecules of the same structural class, preferably the peptide bonds between positions 8/9, 9/10, 15/16, 23/24, 40/41, 53/54, 71/72, 85/86, 94/95, 108/109 and/or 141/142, most preferred between positions 9/10, 40/41, 71/72, 94/95 and/or 108/109 of leptin (SEQ ID NO:127), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of obesity or other diseases connected with leptin.

37. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is a binary toxin.

38. The use according to claim 37, wherein

(i) the engineered protease is capable of hydrolysing peptide bonds in Protective antigen precursor (PA-83) or related molecules of the same structural class, preferably the peptide bonds between positions 72/73, 73/74, 92/93, 93/94, 131/132, 149/150, 178/179, 213/214, 214/215, 387/388, 425/426, 426/427, 427/428, 453/454, 520/521, 608/609, 617/618, 671/672, 679/680, 680/681, 683/684 and/or 684/685, most preferred between positions 72/73, 73/74, 93/94, 149/150, 387/388, 425/426, 427/428 and/or 683/684 of hPA-83 (SEQ ID NO:123), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of anthrax infection or other diseases connected with PA-83.

39. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is a protease, in particular a serine, cysteine or metallo protease.

40. The use according to claim 39, wherein

(i) the protease is capable of hydrolysing peptide bonds in human plasminogen (plasminogen) or related molecules of the same structural class, preferably the peptide bonds between position 580/581 of plasminogen (SEQ ID NO:140), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of thrombosis or other diseases connected with plasminogen.

41. The use according to claim 39, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Prothrombin (thrombin) or related molecules of the same structural

class, preferably the peptide bonds between positions 198/199, 327/328, 363/364, most preferred between positions 327/328 and/or 363/364 of thrombin (SEQ ID NO:149), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of bleeding or other diseases connected with thrombin.

42. The use according to claim 39, wherein

(i) the protease is capable of hydrolysing peptide bonds in human beta-secretase or related molecules of the same structural class, preferably the peptide bonds between positions 61/62, 64/65, 130/131, 131/132, 159/160, 216/217, 238/239, 239/240, 246/247, 256/257, 259/260, 330/331, 365/366, 378/379 and/or 381/382, most preferred between positions 61/62, 131/132, 246/247, 259/260, 365/366 and/or 378/379 of beta-secretase (SEQ ID NO:139), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of Alzheimer or other diseases connected with beta-secretase precursor.

43. The use according to claim 39, wherein

(i) the protease is capable of hydrolysing peptide bonds in human matrix metalloproteinase-2 (hMMP-2) or related molecules of the same structural class, preferably the peptide bonds between positions 62/63, 68/69, 75/76, 76/77, 79/80, 88/89, 110/111, 112/113, 115/116, 120/121, 164/165, 254/255, 267/268, 296/297, 324/325, 325/326, 382/383, 383/384, 470/471, 500/501, 550/551, 564/565, 595/596, 597/598, 608/609, 646/647, 649/650 and/or 650/651, most preferred between positions 68/69, 115/116, 120/121, 164/165, 325/326, 383/384, 470/471, 500/501, 595/596, 608/609 and/or 650/651 of hMMP-2 (SEQ ID NO:131), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of a variety of cancers including bladder cancer, breast tumor cancer, gastric cancer, lung cancer or other diseases connected with hMMP-2.

44. The use according to claim 39, wherein

(i) the protease is capable of hydrolysing peptide bonds in human matrix metalloproteinase-9 (hMMP-9) or related molecules of the same structural class, preferably the peptide bonds between positions 41/42, 42/43, 106/107, 113/114, 134/135, 160/161, 162/163, 163/164, 222/223, 226/227, 265/266, 266/267, 267/268, 284/285, 309/310, 321/322, 322/323, 324/325, 356/357, 380/381, 433/434 and/or 440/441, most preferred between positions 160/161, 163/164, 226/227, 284/285, 321/322, 322/323 and/or 433/434 of hMMP-9 (SEQ ID NO:132), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of a variety of cancers including bladder cancer, breast tumor cancer, gastric cancer, lung cancer or other diseases connected with hMMP-9.

45. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is a type-1 single pass transmembrane protein.

46. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in HIV membrane glycoprotein (GP120) or related molecules of the same structural class, preferably the peptide bonds between positions 97/98, 99/100, 107/108, 113/114, 117/118, 227/228, 231/233, 279/280, 335/336, 337/338, 368/369, 412/413, 419/420, 429/430, 444/445, 457/458, 474/475, 476/477, 477/478, 485/486 and/or 490/491, most preferred between positions 99/100, 368/369, 412/413, 419/420, 444/445 and/or 490/491 of GP120 (SEQ ID NO:124) or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of AIDS or HIV infection or other diseases connected with GP120.

47. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds human Cytotoxic T-lymphocyte protein 4 (CTLA-4) or related molecules of the same structural class, preferably the peptide bonds between positions 14/15, 28/29, 33/34, 38/39, 41/42, 62/63, 72/73, 85/86, 95/96, 100/101, 105/106, 119/120, 125/126 and/or 127/128, most preferred between positions 14/15, 28/29, 38/39, 41/42, 62/63 and/or 85/86 of CTLA-4 (SEQ ID NO:144), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of breast cancer or other diseases connected with CTLA-4.

48. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Integrin alpha-2 (hVLA-2) or related molecules of the same structural class, preferably the peptide bonds between positions 160/161, 174/175, 201/202, 219/220, 231/232, 232/233, 233/234, 243/244, 259/260, 264/265, 268/269, 288/289, 292/293, 294/295, 298/299, 301/302, 310/311 and/or 317/318, most preferred between positions 160/161, 174/175, 201/202, 219/220, 243/244, 264/265, 292/293 and/or 294/295 of hVLA-2 (SEQ ID NO:147), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of renal tumors, uveal melanomas, gastrointestinal tumors or other diseases connected with hVLA-2.

49. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Vascular endothelial growth factor receptor 1 (hVEGFR 1) or related molecules of the same structural class, preferably the peptide bonds between positions 175/176, 180/181, 187/188, 189/190, 190/191, 224/225 and/or 331/332, most preferred between positions 189/190 and/or 331/332 of hVEGFR 1 (SEQ ID NO:114), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of solid tumors and metastatic solid tumors, astrocytic brain tumors, pancreatic cancer, metastatic renal cancer or other diseases connected with hVEGFR 1.

50. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Vascular endothelial growth factor receptor 2 (hVEGFR 2) or related molecules of the same structural class, preferably the peptide bonds between positions 214/215, and/or 323/324, most preferred between positions 214/215 of hVEGFR 2 (SEQ ID NO:115), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of solid tumors and metastatic solid tumors pancreatic cancer, metastatic renal cancer, metastatic CRC, or other diseases connected with hVEGFR 2.

51. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Epidermal growth factor receptor (hEGFr) or related molecules of the same structural class, preferably the peptide bonds between positions 20/21, 29/30, 48/49, 74/75, 155/156, 165/166, 167/168, 202/203, 206/207, 220/221, 223/224, 246/247, 251/252, 254/255, 269/270, 270/271, 297/298, 304/305, 305/306, 357/358, 364/365, 369/370, 430/431, 443/444, 454/455, 455/456, 463/464, 465/466, 476/477, 507/508 and/or 509/510, most preferred between positions 155/156, 246/247, 251/252, 254/255, 269/270, 270/271, 297/298, 304/305, 306/307, 364/365 and/or 454/455 of hEGFr (SEQ ID NO:116), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of bladder cancer, breast cancer, cervical cancer, colorectal cancer, endometrial cancer, oesophageal cancer, head and neck cancer, gastric cancer, non-small-cell lung carcinoma, ovarian cancer or other diseases connected with hEGFr.

52. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Epithelial cell adhesion molecule (Ep-CAM) or related molecules of the same structural class, preferably the peptide bonds between positions 14/15, 19/20, 25/26, 29/30, 30/31, 33/34, 44/45, 55/56, 67/68, 70/71, 90/91 and/or 100/101, most preferred between positions 14/15, 30/31, 44/45, 70/71 and/or 100/101 of Ep-CAM (SEQ ID NO:125) or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of colorectal cancer or other diseases connected with Ep-CAM.

53. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Insulin-like growth factor I receptor (hIGF-1r) or related molecules of the same structural class, preferably the peptide bonds between positions 39/40, 59/60, 115/116, 132/133, 146/147, 171/172, 191/192, 250/251, 262/263, 270/271, 290/291, 306/307, 307/308, 335/336, 336/337, 403/404, 405/406, 455/456 and/or 470/471, most preferred between positions 39/40, 262/263, 306/307, 307/308, 335/336, 405/406 and/or 470/471 of hIGF-1r (SEQ ID NO:126) or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of a variety of cancers including breast cancer or other diseases connected with hIGF-1r.

54. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human T-cell surface antigen CD2 (hCD2) or related molecules of the same structural class, preferably the peptide bonds between positions 16/17, 20/21, 28/29, 29/30, 40/41, 42/43, 43/44, 48/49, 49/50, 51/52, 54/55, 58/59, 63/64, 69/70, 76/77, 89/90 and/or 91/92, most preferred between positions 28/29, 40/41, 43/44, 51/52, 76/77 and/or 89/90 of hCD2 (SEQ ID NO:128) or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of psoriasis or other diseases connected with hCD2.

55. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human T-cell surface glycoprotein CD4 (hCD4) or related molecules of the same structural class, preferably the peptide bonds between positions 53/54, 63/64, 88/89, 166/167, 167/168, 173/174, 206/207, 219/220, 224/225, 226/227, 230/231, 244/245, 251/252, 252/253, 322/323, 329/330 and/or 334/335, most preferred between positions 88/89, 173/174, 206/207, 219/220, 230/231, 251/252 and/or 252/253 of hCD4 (SEQ ID NO:129), or between analogous positions in related molecules;

(ii) preferably the medicament is suitable for the treatment of psoriasis, transplant rejection, graft-versus-host colitis, autoimmune disorders, rheumatoid arthritis or other diseases connected with hCD4.

56. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Integrin alpha-L (hCD11a) or related molecules of the same structural class, preferably the peptide bonds between positions 145/146, 152/153, 156/157, 159/160, 160/161, 177/178, 178/179, 189/190, 190/191, 191/192, 193/194, 197/198, 200/201, 221/222, 229/230, 249/250, 253/254, 268/269, 290/291, 297/298, 304/305 and/or 305/306, most preferred between positions 145/146, 159/160, 160/161, 189/190, 229/230, 249/250, 268/269, 297/298, 304/305 and/or 305/306 of hCD11a (SEQ ID NO:130), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of psoriasis or other diseases connected with hCD11a.

57. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Interferon-gamma receptor alpha chain (hIFN-?-R1) or related

molecules of the same structural class, preferably the peptide bonds between positions 49/50, 52/53, 58/59, 62/63, 72/73, 76/77, 106/107, 107/108, 116/117, 122/123, 174/175, 176/177, 179/180, 215/216 and/or 222/223, most preferred between positions 49/50, 72/73, 116/117, 122/123, 174/175 , 176/177 and/or 215/216 of hIFN-?-R1 (SEQ ID NO:136), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of classical Hodgkin's Lymphoma (cHL), type I diabetes or other diseases connected with hIFN-?-R1.

58. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Platelet membrane glycoprotein IIb/IIIa (GPIIb/IIIa) or related molecules of the same structural class, preferably the peptide bonds between positions 67/68, 76/77, 91/92, 129/130, 143/144, 144/145, 179/180, 181/182, 198/199, 208/209, 209/210, 216/217, 239/240, 261/262, 410/411, 434/435, 532/533, 556/557, 557/558, 596/597, 597/598, 621/622, 650/651, 651/652, 661/662, 662/663 and/or 689/690, most preferred between positions 67/68, 76/77, 179/180, 261/262, 410/411, 434/435, 650/651, 662/663 and/or 689/690 of GPIIb/IIIa (SEQ ID NO:141), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of unstable angina, carotid stenting, ischemic stroke, peripheral vascular diseases, angiogenesis-related diseases, disseminating tumors or other diseases connected with GPIIb/IIIa.

59. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Intercellular adhesion molecule-1 (ICAM-1) or related molecules of the same structural class, preferably the peptide bonds between positions 26/27, 40/41, 60/61, 71/72, 88/89, 97/98, 102/103, 128/129, 131/132, 132/133, 149/150, 150/151, 151/152, 160/161, 164/165 and/or 166/167, most preferred between positions 71/72,

88/89, 102/103, 150/151, 151/152, 160/161 and/or 166/167 of ICAM-1 (SEQ ID NO:142), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of Crohn's disease or other diseases connected with ICAM-1.

60. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds human TGF-beta receptor type II (hTGF-ß RII) or related molecules of the same structural class, preferably the peptide bonds between positions 32/33, 34/35, 35/36, 66/67, 67/68, 69/70, 82/83, 103/104, 104/105, 105/106, 118/119, 122/123 and/or 130/131, most preferred between positions 32/33, 34/35, 66/67, 69/70, 104/105, 122/123 and/or 130/131 of hTGF-ß RII (SEQ ID NO:145), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of diffuse systemic sclerosis or other diseases connected with hTGF-ß RII.

61. The use according to claim 45, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Membrane cofactor protein (hMCP) or related molecules of the same structural class, preferably the peptide bonds between positions 15/16, 17/18, 25/26, 27/28, 31/32, 32/33, 35/36, 47/48, 48/49, 58/59, 67/68, 69/70, 70/71, 110/111, 119/120, 125/126 and/or 130/131, most preferred between positions 15/16, 32/33, 47/48, 48/49, 70/71, 119/120 and/or 125/126 of hMCP (SEQ ID NO:146), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of renal tumors, uveal melanomas, gastrointestinal tumors or other diseases connected with hMCP.

62. The use according to claim 1 or 2, wherein the target substrate hydrolyzed by the protease is a protease-activated receptor or hydroxytryptamine receptor.

63. The use according to claim 62, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Protease activated receptor 1 (hPAR1) or related molecules of the same structural class, preferably the peptide bonds between positions 46/47, 50/51, 51/52, 52/53 and/or 58/59 of hPAR1 (SEQ ID NO:110), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of thrombosis or other diseases connected with hPAR1.

64. The use according to claim 62, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Protease activated receptor 2 (hPAR2) or related molecules of the same structural class, preferably the peptide bonds between positions 41/42, 44/45, 51/52, 59/60 and/or 62/63 of hPAR2 (SEQ ID NO:111), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of Crohn's disease, Ulcerative colitis and Inflammatory bowel disease, asthma, inflammation associated pain and arthritis or other diseases connected with hPAR2.

65. The use according to claim 62, wherein

(i) the protease is capable of hydrolysing peptide bonds in human Protease activated receptor 4 (hPAR4) or related molecules of the same structural class, preferably the peptide bonds between positions 57/58, 59/60, 68/69, 74/75 and/or 78/79 of hPAR4 (SEQ ID NO:113), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of thrombosis or other diseases connected with hPAR4.

66. The use according to claim 62, wherein

(i) the protease is capable of hydrolysing peptide bonds in human 5-hydroxytryptamine 1A receptor (h5-HT-1A) or related molecules of the same structural class, preferably the peptide bonds between

positions 101/102, 102/103, 181/182 and/or 370/371 of h5-HT-1A (SEQ ID NO:117) or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of irritable bowel syndrome or other diseases connected with h5-HT-1A.

67. The use according to claim 1 or 2, wherein

(i) the protease is capable of hydrolysing peptide bonds in human carcinoembryonic antigen (hCEA) or related molecules of the same structural class, preferably the peptide bonds between positions 17/18, 69/70, 71/72, 74/75, 77/78, 98/99, 116/117, 126/127 and/or 128/129 of hCEA (SEQ ID NO:138), or between analogous positions in related molecules; and/or

(ii) the medicament is suitable for the treatment of colon cancer or other diseases connected with hCEA.

68. The use according to any one of claims 1 to 67, wherein the protease is an engineered protease,preferably an engineered protease characterized by a combination of the following components:
(a) a protein scaffold capable to catalyze at least one chemical reaction on at least one target substrate and being derived from one or more proteins, and
(b) one or more specificity determining regions (SDRs) located at sites in the protein scaffold that enable the resulting engineered protein to discriminate between at least one target substrate and one or more different substrates, and wherein the SDRs are essentially synthetic peptide sequences.

69. The use according to claim 68, wherein
(I) the SDRs (b) have a length between one and 50 amino acid residues, preferably have a length between 2 and 20 amino acid residues, more preferably a length between 2 and 10 amino acid residues, even more preferably a length between 3 and 8 amino acid residues, and wherein the number of SDRs is at least one, preferably more than one, more preferably between two and eleven, most preferably between two and six; and/or
(II) the protein scaffold (a) is comprised of one or more polypeptide segments being derived from same or different

(i) proteins encoded by a gene of viral or prokaryotic or eukaryotic origin, and/or
(ii) native enzymes, mutated variants or truncated derivates thereof, and/or
(iii) mammalian enzymes, preferably human enzymes.

70. The use according to claim 68 or 69, wherein the protein scaffold (a) is derived from a protease selected from the group consisting of aspartic, cysteine, serine, metallo and threonine proteases,
even more preferably the protein scaffold (a) is derived from a serine protease of the structural class S1, S8, S11, S21, S26, S33 or S51, most preferably from class S1 or S8, a cysteine protease of the structure class C1, C2, C4, C10, C14, C19, C47, C48 or C56, most preferably from class C14, or an aspartic protease of the structural class A1, A2 or A26, most preferably from class A1, or a metalloprotease of the strucutral class M4 or M10.

71. The use according to any one of claims 68 to 70, wherein

(i) the protein scaffold (a) is derived from a serine protease of the structural class S1; and/or

(ii) the SDRs are located at one or more positions selected from the group of positions that correspond structurally or by amino acid sequence homology to the regions 18-25, 38-48, 54-63, 73-86, 122-130, 148-156, 165-171 and 194-204 in human trypsin I having the amino acid sequence shown in SEQ ID NO:1, and preferably at one or more positions selected from the group of positions that correspond structurally or by amino acid sequence homology to the regions 20-23, 41-45, 57-60, 76-83, 125-128, 150-153, 167-169 and 197-201 in human trypsin I.

72. The use according to claim 71, wherein

(i) the protein scaffold (a) is derived from the serine protease trypsin, preferably human trypsin I having the amino acid sequence shown in SEQ ID NO:1, or the amino acid sequence SEQ ID NO:1 comprising one or more of the following amino acid substitutions E56G, R78W, Y131F, A146T and C183R; and

(ii) at least one of two SDRs are located in the scaffold, a first SDR having a length of up to 6 amino acids and bein inserted between

residues 42 and 43, and a second SDR having a length of up to 5 amino acids and bein inserted between residues 123 and 124 (numbering relative to human trypsin having the amino acid sequence shown in SEQ ID NO:1).

73. The use according to claim 72, wherein

(i) one of the peptide sequences of the following group: SEQ ID NO: 72, 78, 79, 80, 84, 85, 86, 87, 88, and 89 is inserted as the first SDR between residues 42 and 43; and/or one of the peptide sequences of the following group: SEQ ID NO: 73, 81, 82, 83, 90, 91, 92, 93, 94, and 95 is inserted as the second SDR between residues 123 and 124; or

(ii) the engineered enzyme comprises an amino acid sequence as shown in SEQ ID NO: 74, or SEQ ID NO: 75.

74. The use according to any one of claims 68 to 70, wherein

(i) the protein scaffold (a) is derived from a serine protease of the structural class S8; and/or

(ii) (ii) the SDRs are located at one or more positions selected from the group of positions that correspond structurally or by amino acid sequence homology to the regions 6-17, 25-29, 47-55, 59-69, 101-111, 117-125, 129-137, 139-154, 158-169, 185-195 and 204-225 in subtilisin E from *Bacillus subtilis* having the amino acid shown in SEQ ID NO:7, and preferably at one or more positions selected from the group of positions that correspond structurally or by amino acid sequence homology to the regions 59-69, 101-111, 129-137, 158-169 and 204-225in subtilisin E for *Bacillus subtilis*.

75. The use according to any one of claims 68 to 70, wherein

(i) the protein scaffold (a) is derived from an aspartic protease of the structural class A1; and/or

(ii) the SDRs are located at one or more positions selected from the group of positions that correspond structurally or by amino acid sequence homology to the regions 6-18, 49-55, 74-83, 91-97, 112-120, 126-137, 159-164, 184-194, 242-247, 262-267 and 277-300

in human pepsin having the amino acid sequence shown in SEQ ID NO:11, and more preferably at one or more positions selected from the group of positions that correspond structurally or by amino acid sequence homology to the regions 10-15, 75-80, 114-118, 130-134, 186-191 and 280-296 in human pepsin.

76. The use according to any one of claims 68 to 70, wherein

(i) the protein scaffold (a) is derived from a cysteine protease of the structural class C14; and/or

(ii) the SDRs are located at one or more positions selected from the group of positions that correspond structurally or by amino acid sequence homology to the regions 78-91, 144-160, 186-198, 226-243 and 271-291 in human caspase 7 having the amino acid sequence of SEQ ID NO:14, and preferably at one or more positions selected from the group of positions that correspond structurally or by amino acid sequence homology to the regions 80-86, 149-157, 190-194 and 233-238 of human caspase 7.

77. The use according to any one of claims 1 to 76, wherein the protease comprises

(i) at least one further proteinacious component, preferably being selected from the group consisting of binding domains, receptors, antibodies, regulation domains, pro-sequences, and fragments thereof, and/or

(ii) at least one further functional component, preferably being selected from the group consisting of polyethylenglycols, carbohydrates, lipids, fatty acids, nucleic acids, metals, metal chelates, and fragments or derivatives thereof.

78. The use according to any one of claims 1 to 77, wherein the protease is obtainable by a method comprising at least the following steps:

(a) providing a protein scaffold which catalyzes at least one chemical reaction on at least one target substrate,

(b) generating a library of enzymes or isolated enzymes by combining the protein scaffold from step (a) with variants of one or more fully or partially random peptide sequences at sites in the protein scaffold that enable the

resulting enzyme to discriminate between at least one target substrate and one or more different substrates, and
(c) selecting out of the (library of) enzymes generated in step (b) one or more enzymes that have defined specificities towards at least one target substrate.

79. Use of an enzyme according to any one of claims 1 to 78 for *in vivo* or *in vitro* diagnostic purposes.

80. A pharmaceutical or diagnostic composition comprising one or more enzymes according to any one of claims 1 to 78, said pharmaceutical or diagnostic composition optionally comprising pharmaceutically or diagnostically acceptable carrier(s), excipient(s) and/or auxiliary agent(s).

81. A method for cleaving a target substrate as defined in claims 1 to 67 *in vivo* or *in vitro*, which comprises contacting the target substrate with a protease as defined in claims 1 to 78.

81. A method for treatment of a disease in a patient connected with a specific target substrate as defined in anyone of claims 1 to 67, which comprises administering the patent a suitable amount of a protease with defined specificity for said specific target substrate as defined in anyone of claims 1 to 78.

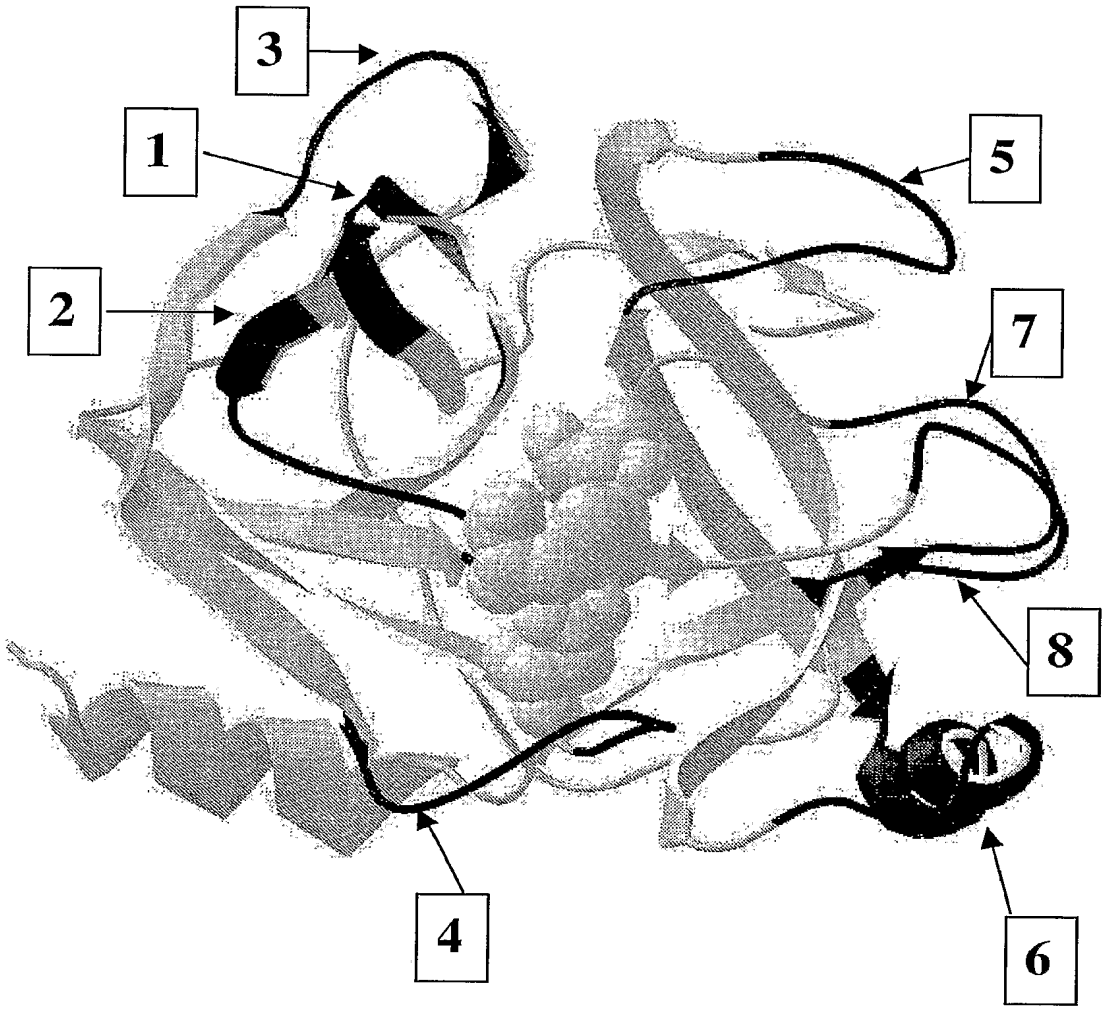

Fig. 1

```
Trypsin           IVGGYNCEENSVPYQVSL---NSGYHF-CGGSLINEQWVVSAGHCY----
a-Thrombin        IVEGSDAEIGMSPWQVMLFRKSPQELL-CGASLISDRWVLTAAHCLLYPP
Enteropeptidase   IVGGSNAKEGAWPWVVGL---YYGGRLLCGASLVSSDWLVSAAHCVYGRN
                  ** *        *  * *          ** **    *   * ** ____

Trypsin           ------KSRIQVRLGEH---NIEVLEGN-EQFINAAKIIRHPQYD-RKTL
a-Thrombin        WDKNFTENDLLVRIGKH---SRTRYERNIEKISMLEKIYIHPRYNWRENL
Enteropeptidase   LE----PSKWTAILGLHMKSNLTSPQTV-PRLID--EIVINPHYN-RRRK
                  -1----        * *                      *   * *  *

Trypsin           NNDIMLIKLSSRAVINARVSTISLPTA----PPAT-----GTKCLISGWG
a-Thrombin        DRDIALMKLKKPVAFSDYIHPVCLPDR----ETAASLLQAGYKGRVTGWG
Enteropeptidase   DNDIAMMHLEFKVNYTDYIQPICLPEENQVFPP-------GRNCSIAGWG
                    **    *              **  ------2------*      ***

Trypsin           N-----TASSGADYPDELQCLDAPVLSQAKCEASYPG-KITSNMFCVGFL
a-Thrombin        NLKETWTANVGKGQPSVLQVVNLPIVERPVCKDSTRI-RITDNMFCAGYK
Enteropeptidase   T-----VVYQGTT-ANILQEADVPLLSNERCQQQMPEYNITENMICAGYE
                   --3--    *      **    *       *         ** ** * *

Trypsin           -EGGK--DSCQGDSGGPVVCNGQ----LQ------GVVSWGDGCAQKNKP
a-Thrombin        PDEGKRGDACEGDSGGPFVMKSP----FNNRWYQMGIVSWGEGCDRDGKY
Enteropeptidase   -EGGI--DSCQGDSGGPLMCQENNRWFLA------GVTSFGYKCALPNRP
                     *   * * ******      ------4-----*  * *  *

Trypsin           GVYTKVYNYVKWIKNTIAANS-
a-Thrombin        GFYTHVFRLKKWIQKVIDQFGE
Enteropeptidase   GVYARVSRFTEWIQSFLH----
                  * *  *     **
```

Fig. 2

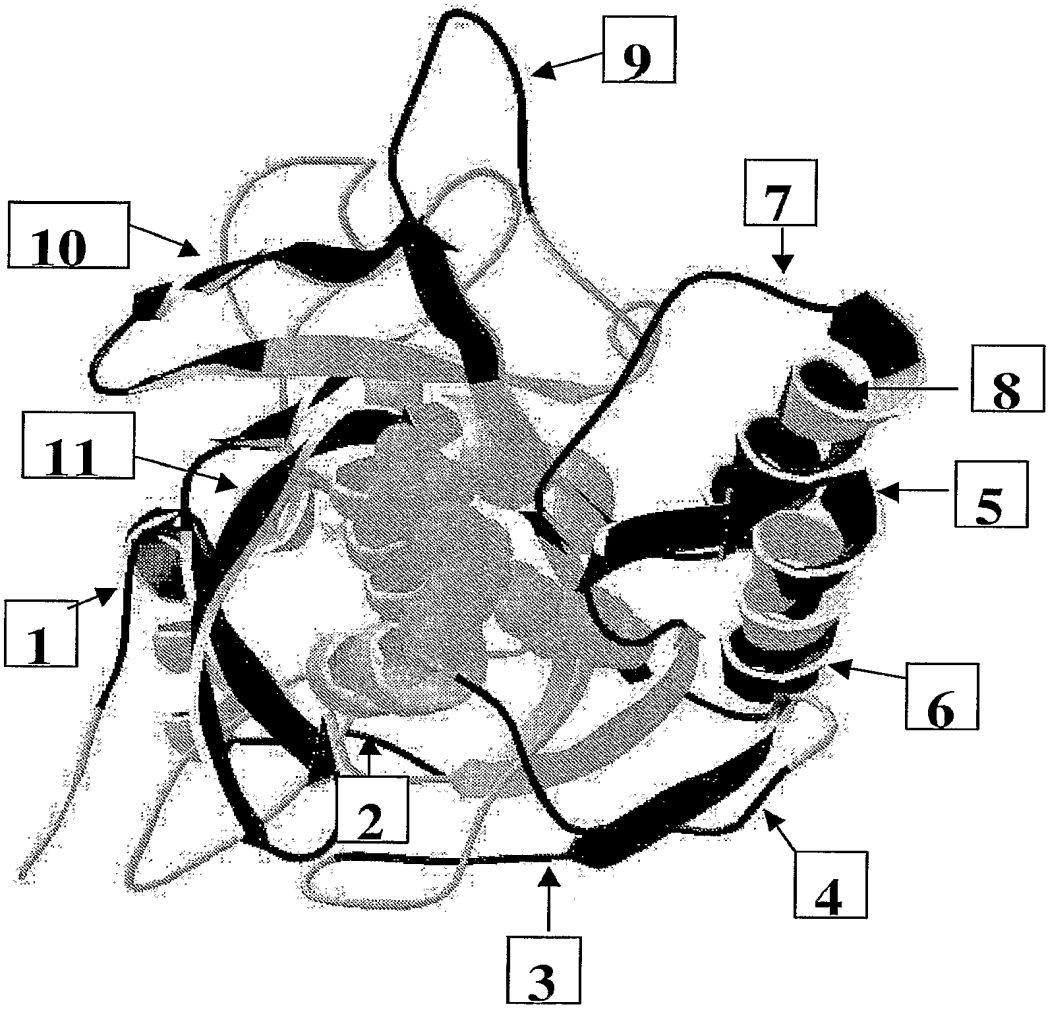

Fig. 3

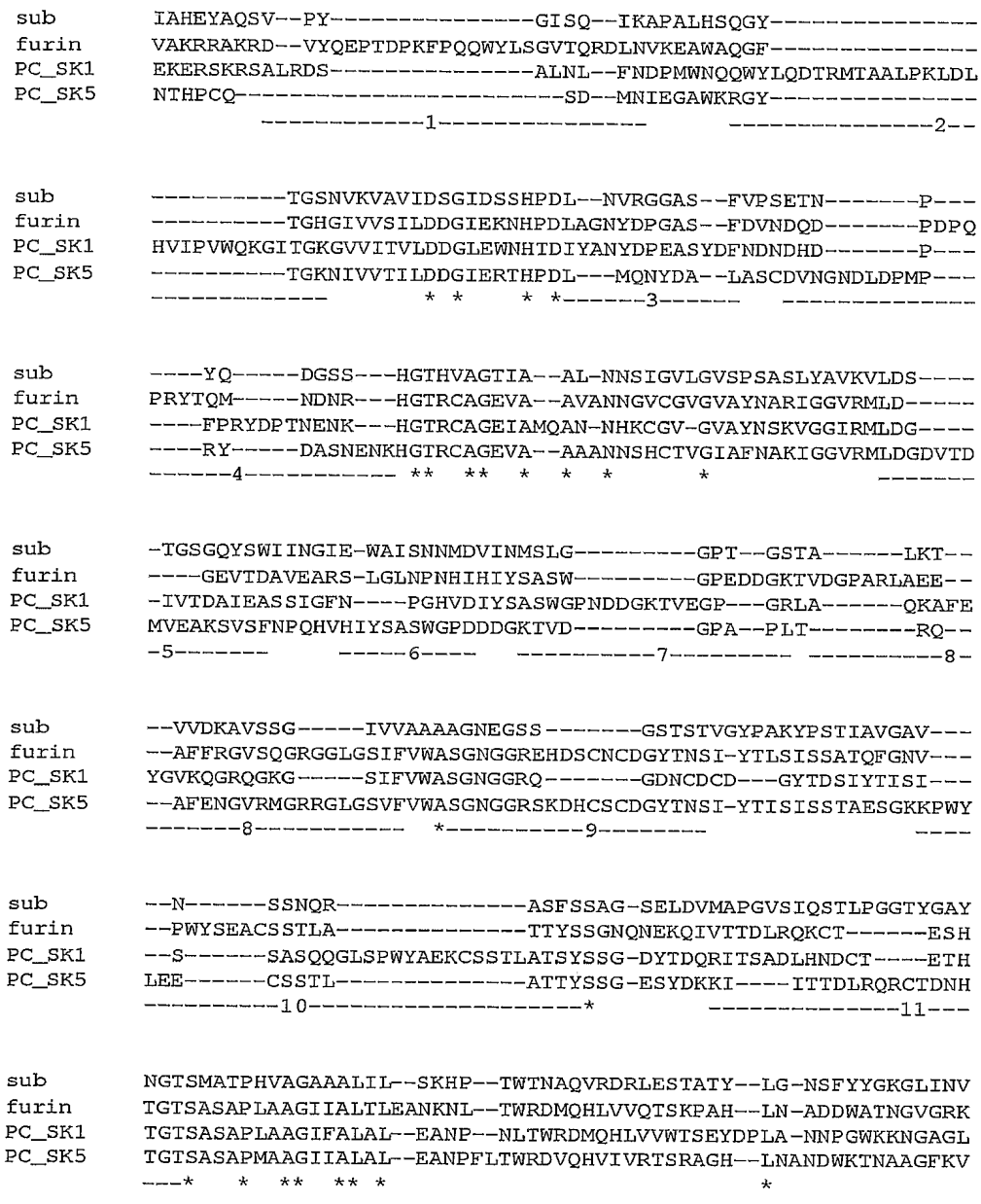
sub IAHEYAQSV--PY----------GISQ--IKAPALHSQGY----------
furin VAKRRAKRD--VYQEPTDPKFPQQWYLSGVTQRDLNVKEAWAQGF----------
PC_SK1 EKERSKRSALRDS----------ALNL--FNDPMWNQQWYLQDTRMTAALPKLDL
PC_SK5 NTHPCQ----------SD--MNIEGAWKRGY----------
1
2
sub ----------TGSNVKVAVIDSGIDSSHPDL--NVRGGAS--FVPSETN-------P---
furin ----------TGHGIVVSILDDGIEKNHPDLAGNYDPGAS--FDVNDQD--------PDPQ
PC_SK1 HVIPVWQKGITGKGVVITVLDDGLEWNHIDIYANYDPEASYDFNDNDHD-------P---
PC_SK5 ----------TGKNIVVTILDDGIERTHPDL---MQNYDA--LASCDVNGNDLDPMP---
3
sub ----YQ-----DGSS---HGTHVAGTIA--AL-NNSIGVLGVSPSASLYAVKVLDS----
furin PRYTQM-----NDNR---HGTRCAGEVA--AVANNGVCGVGVAYNARIGGVRMLD-----
PC_SK1 ----FPRYDPTNENK---HGTRCAGEIAMQAN-NHKCGV-GVAYNSKVGGIRMLDG----
PC_SK5 ----RY-----DASNENKHGTRCAGEVA--AAANNSHCTVGIAFNAKIGGVRMLDGDVTD
4
sub -TGSGQYSWIINGIE-WAISNNMDVINMSLG---------GPT--GSTA------LKT--
furin ----GEVTDAVEARS-LGLNPNHIHIYSASW---------GPEDDGKTVDGPARLAEE--
PC_SK1 -IVTDAIEASSIGFN----PGHVDIYSASWGPNDDGKTVEGP---GRLA------QKAFE
PC_SK5 MVEAKSVSFNPQHVHIYSASWGPDDDGKTVD---------GPA--PLT--------RQ--
5
6
7
8
sub --VVDKAVSSG-----IVVAAAAGNEGSS-------GSTSTVGYPAKYPSTIAVGAV---
furin --AFFRGVSQGRGGLGSIFVWASGNGGREHDSCNCDGYTNSI-YTLSISSATQFGNV---
PC_SK1 YGVKQGRQGKG-----SIFVWASGNGGRQ-------GDNCDCD---GYTDSIYTISI---
PC_SK5 --AFENGVRMGRRGLGSVFVWASGNGGRSKDHCSCDGYTNSI-YTISISSTAESGKKPWY
8
9
sub --N------SSNQR----------ASFSSAG-SELDVMAPGVSIQSTLPGGTYGAY
furin --PWYSEACSSTLA----------TTYSSGNQNEKQIVTTDLRQKCT------ESH
PC_SK1 --S------SASQQGLSPWYAEKCSSTLATSYSSG-DYTDQRITSADLHNDCT----ETH
PC_SK5 LEE------CSSTL----------ATTYSSG-ESYDKKI----ITTDLRQRCTDNH
10
11
sub NGTSMATPHVAGAAALIL--SKHP--TWTNAQVRDRLESTATY--LG-NSFYYGKGLINV
furin TGTSASAPLAAGIIALTLEANKNL--TWRDMQHLVVQTSKPAH--LN-ADDWATNGVGRK
PC_SK1 TGTSASAPLAAGIFALAL--EANP--NLTWRDMQHLVVWTSEYDPLA-NNPGWKKNGAGL
PC_SK5 TGTSASAPMAAGIIALAL--EANPFLTWRDVQHVIVRTSRAGH--LNANDWKTNAAGFKV

Fig. 4

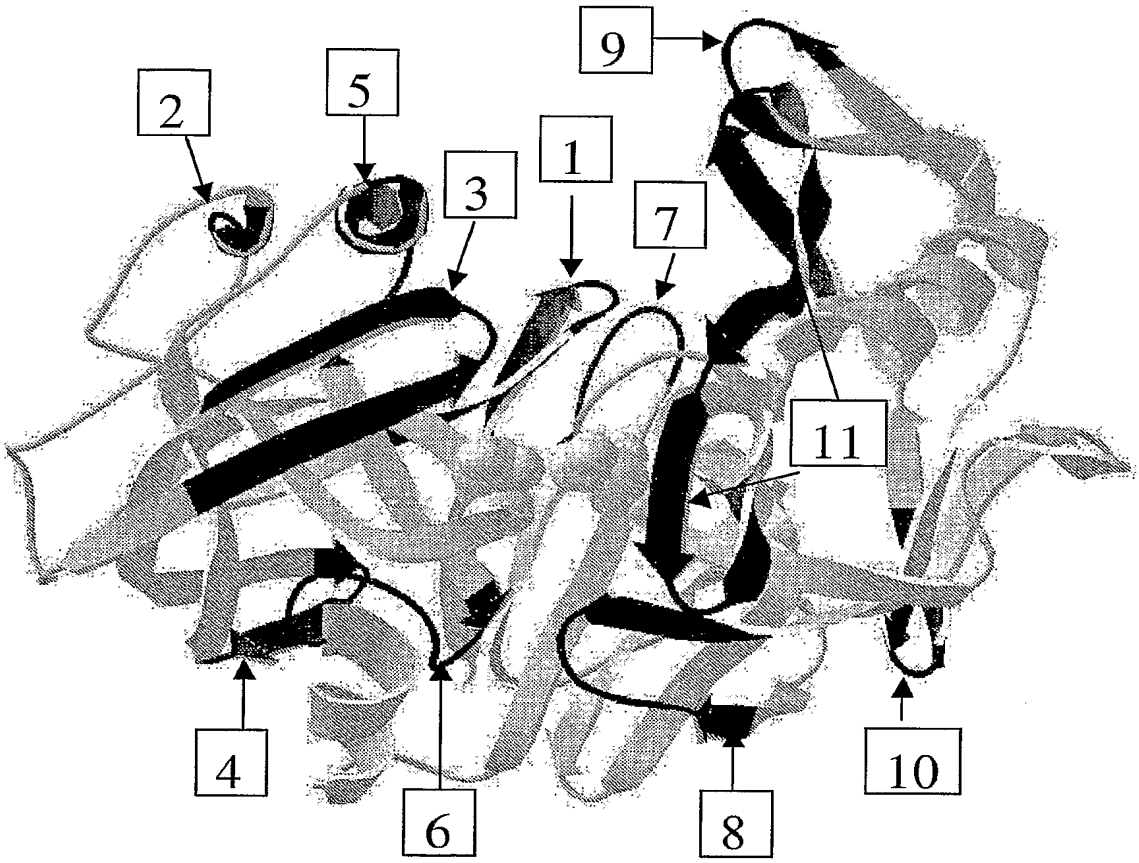

Fig.5

```
Peps.   TLVDEQP----LENYLDMEYFGTIGIGTPAQDFTVVFDTGSSNLWVPSVYCSSL--ACTN
Secr.   EMVDN------LRGKSGQGYYVEMTVGSPPQTLNILVDTGSSNFAVGAAPHPFL------
Cath.   PAVTEGPIPEVLKNYMDAQYYGEIGIGTPPQCFTVVFDTGSSNLWVPSIHCKLLDIACWI
          * ----1-----     *      * * *      ******  *      -----2--

Peps.   HNRFNPEDSSTYQSTSETVSITYGTGSMTGILGYDTVQV----------G---GISDTN
Secr.   HRYYQRQLSSTYRDLRKGVYVPYTQGKWEGELGTDLVSI----------PHGPNVTVRA
Cath.   HHKYNSDKSSTYVKNGTSFDIHYGSGSLSGYLSQDTVSVPCQSASSASALG---GVKVER
        -       ****       -----3----* *  * ---------4----------

Peps.   QIFGLSETEPGSFLYYAPFDGILGLAYPSIS--SSGATPVFDNIWNQGLVSQDLFSVYLS
Secr.   NIAAITESDK-FFINGSNWEGILGLAYAEIARPDDSLEPFFDSLVKQTHVP-NLFSLQLC
Cath.   QVFGEATKQPGITFIAAKFDGILGMAYPRIS--VNNVLPVFDNLMQQKLVDQNIFSFYLS
                ----5----**** ------6------- **    *  *    **  *

Peps.   ADD----------KS--GSVVIFGGIDSSYYTGSLNWVPVTVEGYWQITVDSITMNGETI
Secr.   GAGFPLNQSEVLASV--GGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDL
Cath.   RDP----------DAQPGGELMLGGTDSKYYKGSLSYLNVTRKAYWQVHLDQVEVASGLT
        --------7---------     ** *   * ***  -----8-----

Peps.   A--CAEGC--QAIVDTGTSLLTGPTSPIANIQSDIGASENSD--------GDMVVSCSAI
Secr.   KMDCKEYNYDKSIVDSGTTNLRLPKKVFEAAVKSIKAASSTEKFPDGFWLGEQLV-CWQA
Cath.   L--CKEGC--EAIVDTGTSLMVGPVDEVRELQKAIGAVPLIQ--------GEYMIPCEKV
          * *      *** **     *          * *  -------9------ *

Peps.   SSLPDIVFTI----------NGVQYPVPPSAYILQSEGS-----CISGFQGMNVP-TESG
Secr.   GTTPWNIFPVISLYLMGEVTNQSFRITILPQQYLRPVEDV----ATSQDDCYKFAISQSS
Cath.   STLPAITLKL----------GGKGYKLSPEDYTLKVSQAGKTLCLSGFMGMDIP-PPSG
          *   -------10-------    *  * ---------------11----------

Peps.   ELWILGDVFIRQYFTVFDRANNQVGLAPVA
Secr.   TGTVMGAVIMEGFYVVFDRARKRIGFAVSA
Cath.   PLWILGDVFIGRYYTVFDRDNNRVGFAEAA
        --   * *       ****     * *  *
```

Fig. 6

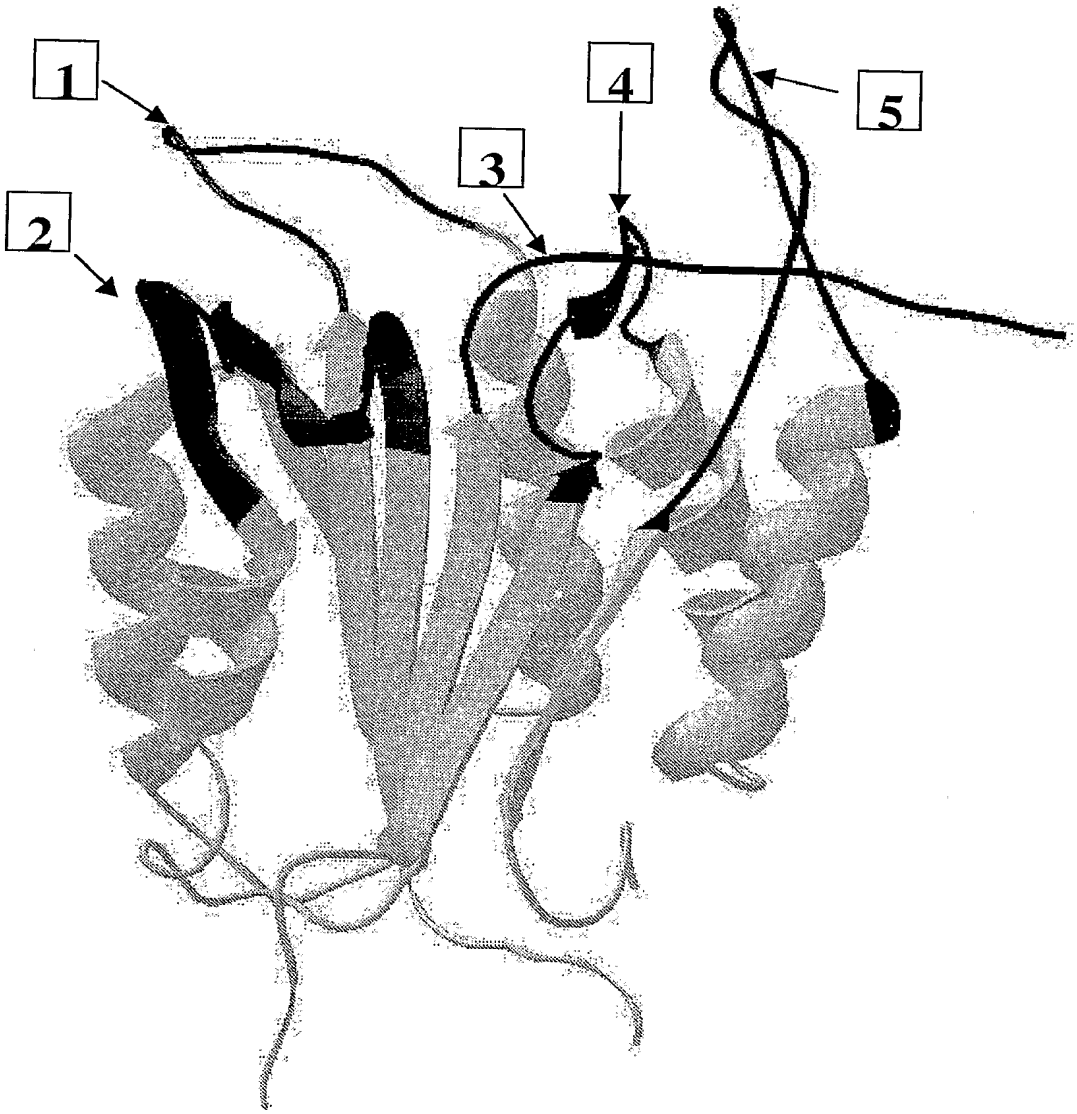

Fig. 7

```
01  MLEADDQGCI EEQGVEDSAN EDSVDAKPDR SSFVPSLFSK KKKNVTMRSI KTTRDRVPTY

61  QYNMNFEKLG KCIIINNKNF DKVTGMGVRN GTDKDAEALF KCFRSLGFDV IVYNDCSCAK
                          ------1---------

121 MQDLLKKASE EDHTNAACFA CILLSHGEEN VIYGKDGVTP IKDLTAHFRG DRSKTLLEKP
                               --------2---------

181 KLFFIQACRG TELDDGIQAD SGPINDTDAN PRYKIPVEAD FLFAYSTVPG YYSWRSPGRG
      -----3--------                                  -----------4----

241 SWFVQALCSI LEEHGKDLEI MQILTRVNDR VARHFESQSD DPHFHEKKQI PCVVSMLTKE
    ---                                 ----------5------------

301 LYFSQ
```

Fig. 8

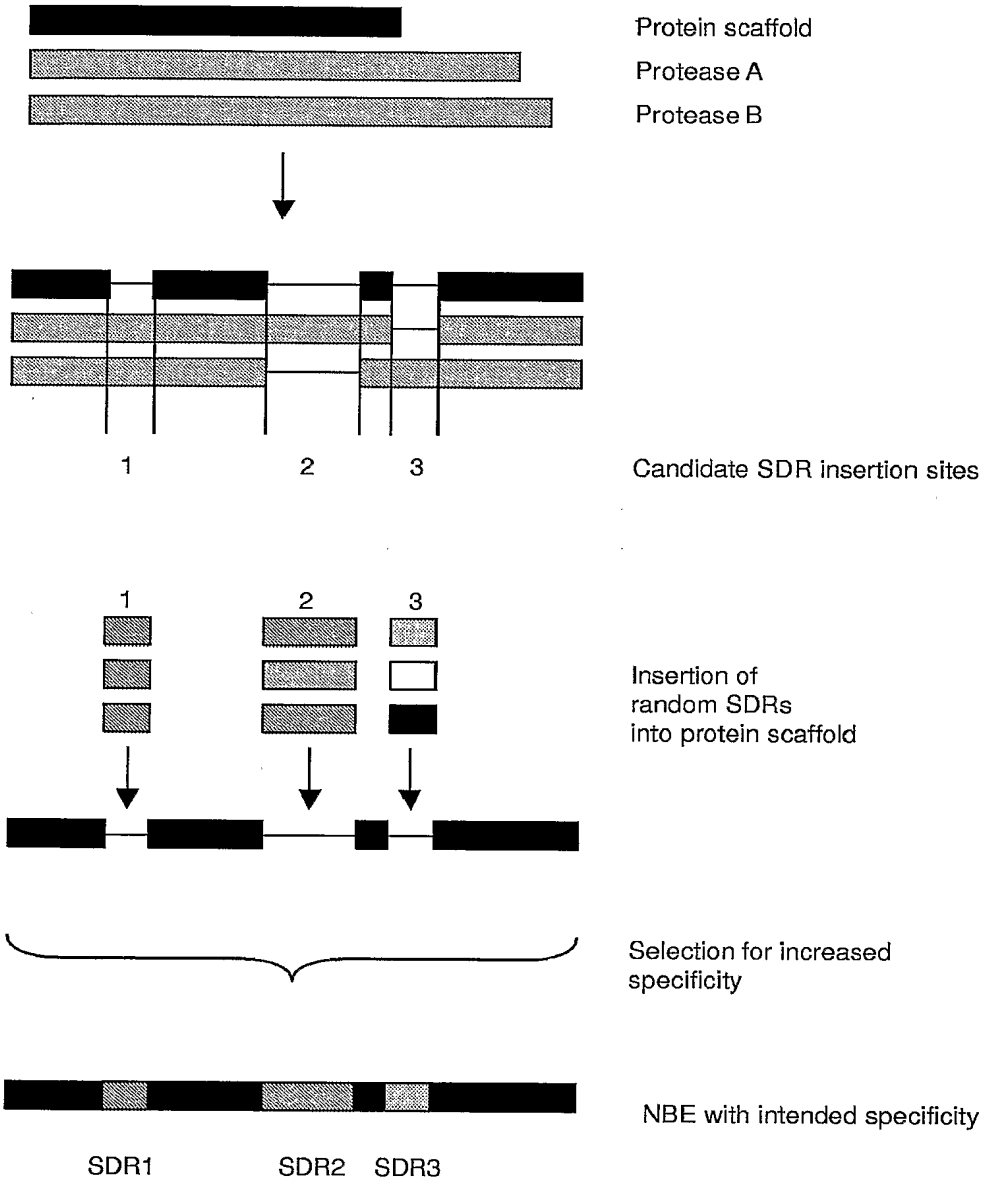
Protein scaffold
Protease A
Protease B
1
2
3
Candidate SDR insertion sites
1
2
3
Insertion of
random SDRs
into protein scaffold
Selection for increased
specificity
NBE with intended specificity
SDR1
SDR2
SDR3

Fig. 9

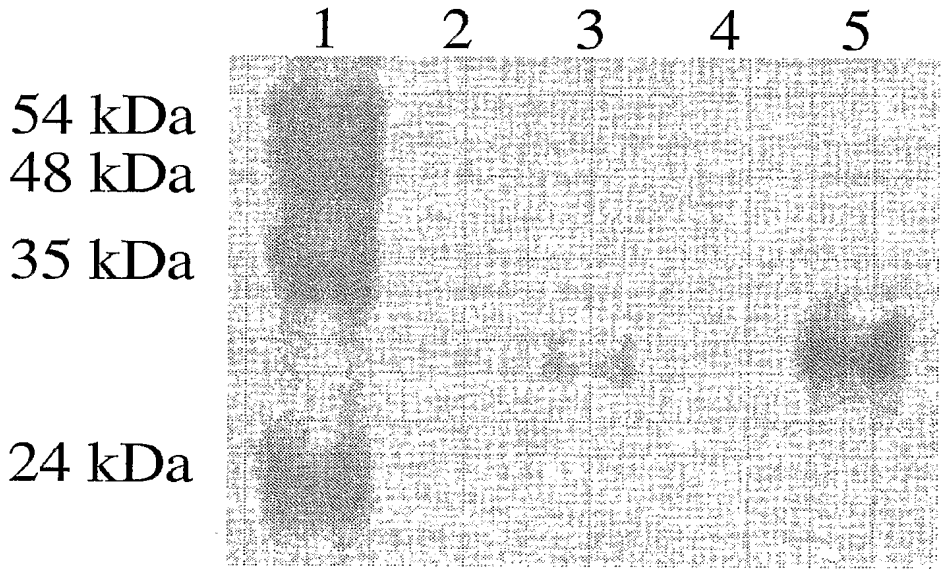

Fig. 10

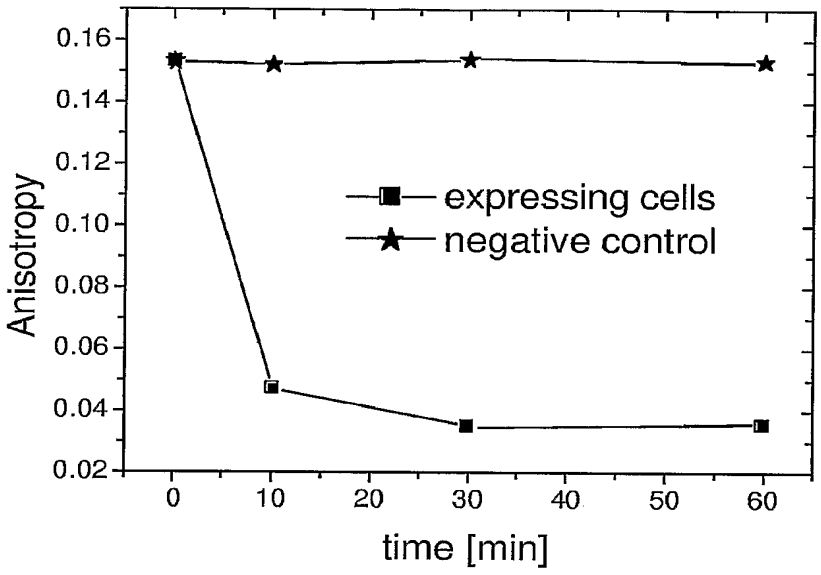
Anisotropy
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0
10
20
30
40
50
60
time [min]
expressing cells
negative control

Fig. 11

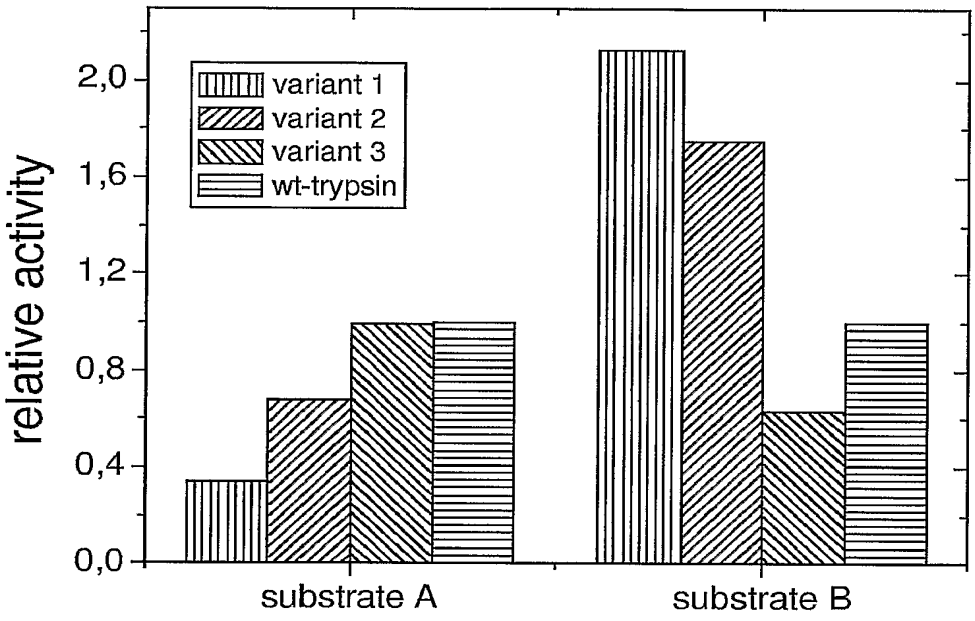
relative activity
2,0
1,6
1,2
0,8
0,4
0,0
variant 1
variant 2
variant 3
wt-trypsin
substrate A
substrate B

Fig. 12

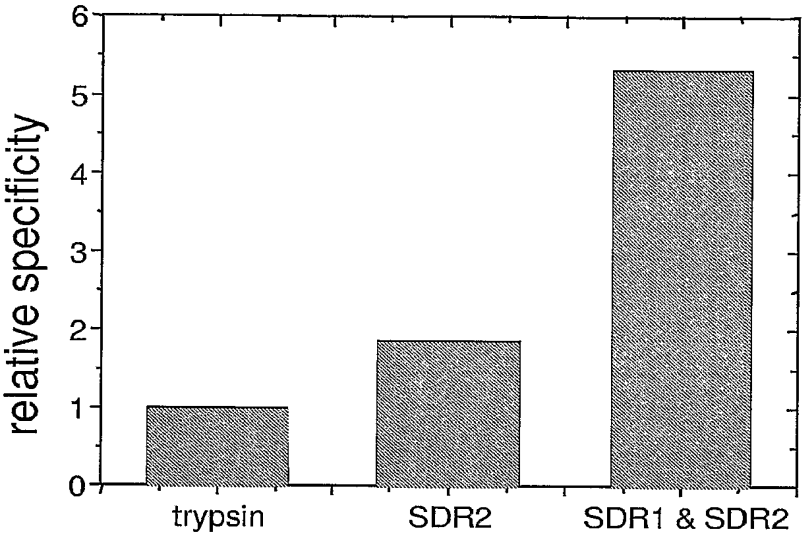

Fig. 13

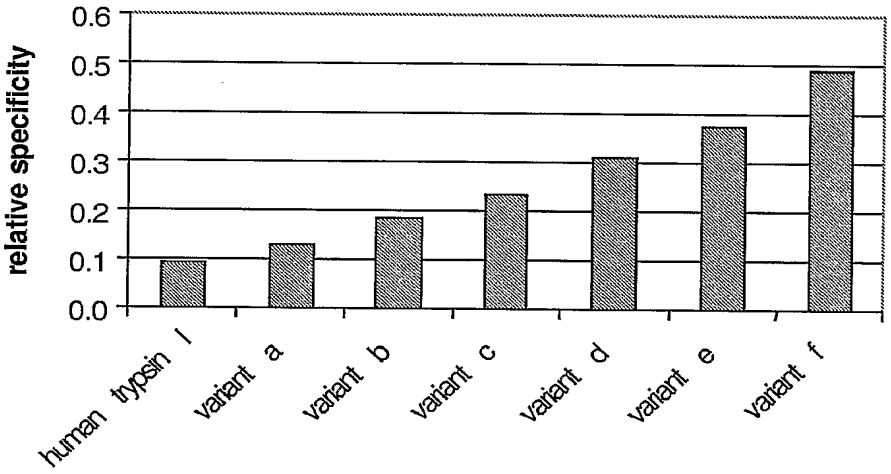
relative specificity
0.6
0.5
0.4
0.3
0.2
0.1
0.0
human trypsin I
variant a
variant b
variant c
variant d
variant e
variant f

Fig. 14

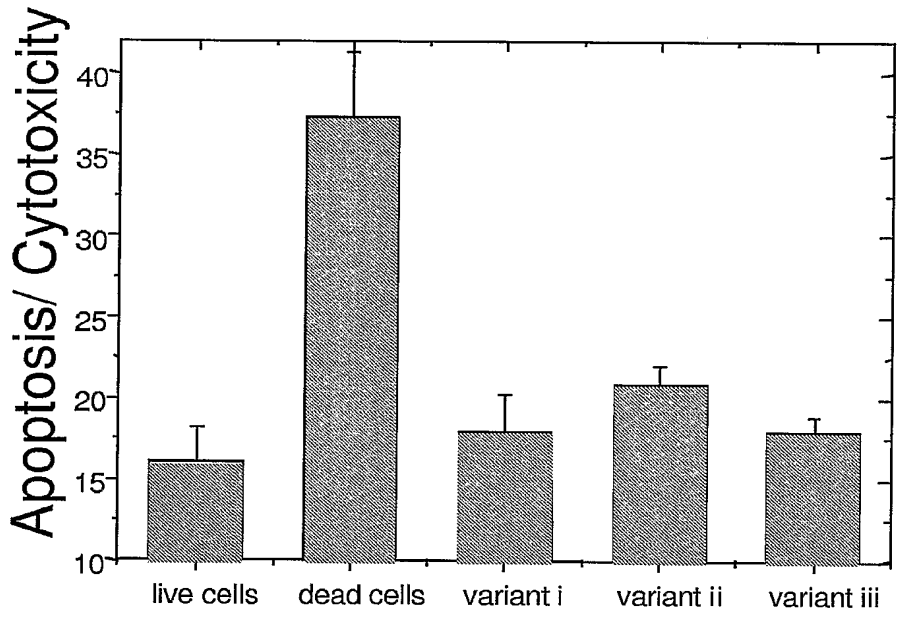
Apoptosis/ Cytotoxicity
40
35
30
25
20
15
10
live cells
dead cells
variant i
variant ii
variant iii

Fig. 15

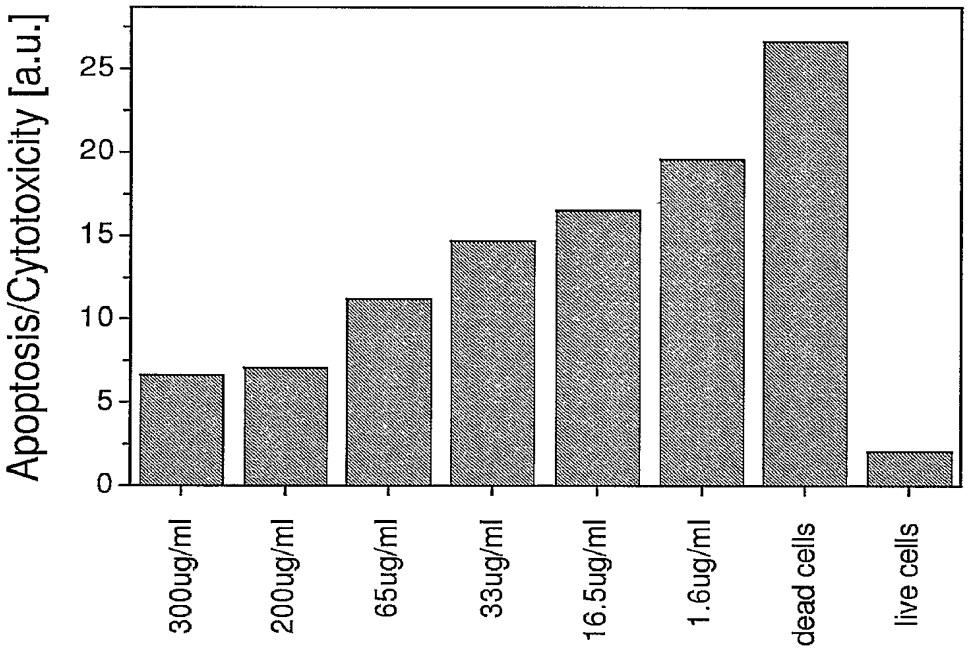
Apoptosis/Cytotoxicity [a.u.]
0
5
10
15
20
25
300ug/ml
200ug/ml
65ug/ml
33ug/ml
16.5ug/ml
1.6ug/ml
dead cells
live cells

Fig. 16

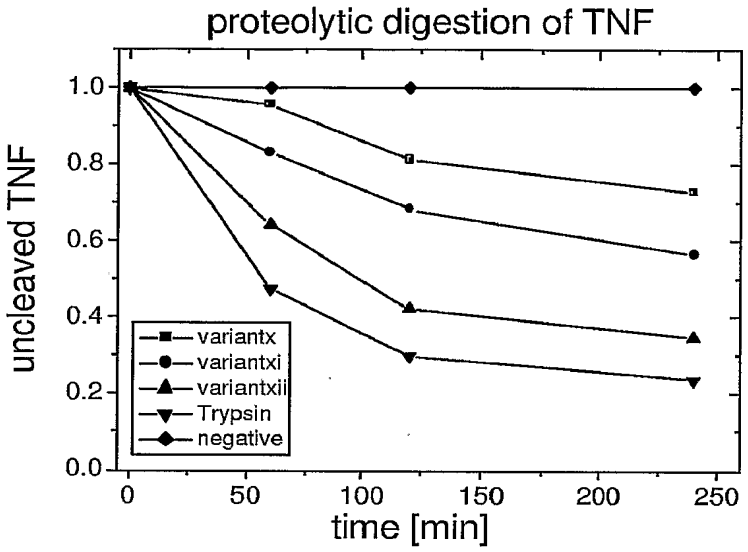

Fig. 17a

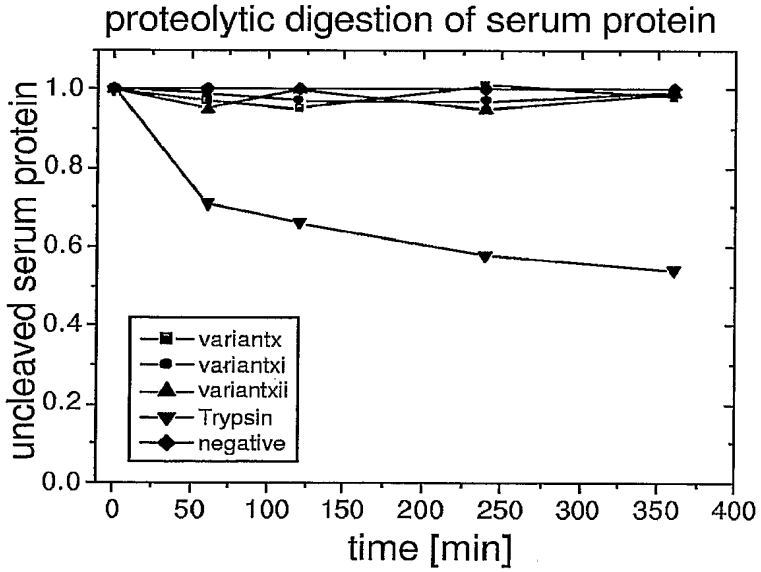

Fig. 17b

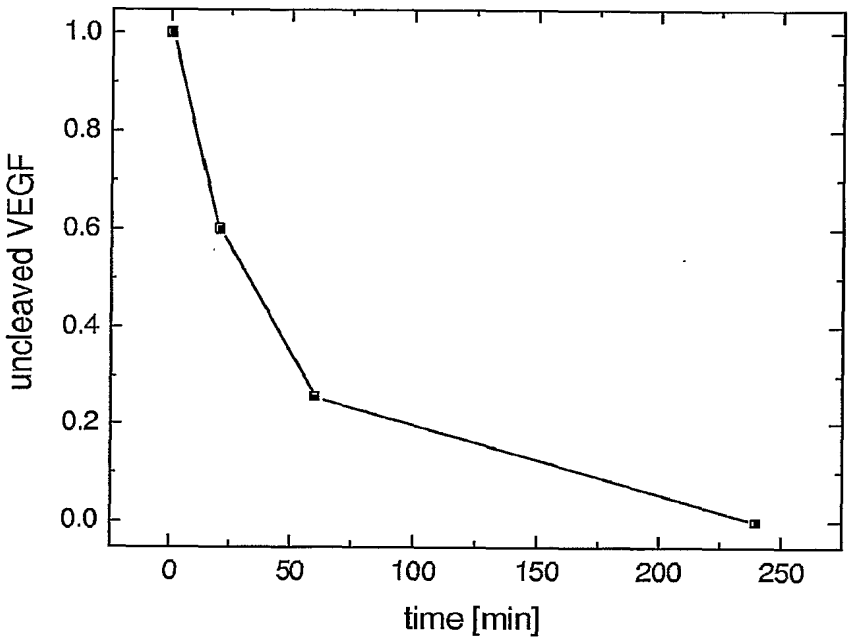

Fig. 18

SEQUENCE LISTING

<110> DIREVO Biotech AG
<120> NEW BIOLOGICAL ENTITIES AND THE USE THEREOF
<130> 041481wo/JH/ml
<160> 126
<170> PatentIn version 3.1

<210> 1
<211> 224
<212> PRT
<213> Homo sapiens

<400> 1
Ile Val Gly Gly Tyr Asn Cys Glu Glu Asn Ser Val Pro Tyr Gln Val
1 5 10 15
Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Glu
20 25 30
Gln Trp Val Val Ser Ala Gly His Cys Tyr Lys Ser Arg Ile Gln Val
35 40 45
Arg Leu Gly Glu His Asn Ile Glu Val Leu Glu Gly Asn Glu Gln Phe
50 55 60
Ile Asn Ala Ala Lys Ile Ile Arg His Pro Gln Tyr Asp Arg Lys Thr
65 70 75 80
Leu Asn Asn Asp Ile Met Leu Ile Lys Leu Ser Ser Arg Ala Val Ile
85 90 95
Asn Ala Arg Val Ser Thr Ile Ser Leu Pro Thr Ala Pro Pro Ala Thr
100 105 110
Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly Asn Thr Ala Ser Ser Gly
115 120 125
Ala Asp Tyr Pro Asp Glu Leu Gln Cys Leu Asp Ala Pro Val Leu Ser
130 135 140
Gln Ala Lys Cys Glu Ala Ser Tyr Pro Gly Lys Ile Thr Ser Asn Met
145 150 155 160
Phe Cys Val Gly Phe Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp
165 170 175
Ser Gly Gly Pro Val Val Cys Asn Gly Gln Leu Gln Gly Val Val Ser
180 185 190
Trp Gly Asp Gly Cys Ala Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys
195 200 205
Val Tyr Asn Tyr Val Lys Trp Ile Lys Asn Thr Ile Ala Ala Asn Ser
210 215 220

<210> 2
<211> 235
<212> PRT
<213> artificial sequence
<220>
<223> trypsin variant 1

<400> 2
Ile Val Gly Gly Tyr Asn Cys Glu Glu Asn Ser Val Pro Tyr Gln Val
1 5 10 15
Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Glu
20 25 30
Gln Trp Val Val Ser Ala Gly His Cys Tyr Asp Ala Val Gly Arg Asp
35 40 45
Lys Ser Arg Ile Gln Val Arg Leu Gly Glu His Asn Ile Glu Val Leu
50 55 60

```
Glu Gly Asn Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro
65                  70                  75                  80
Gln Tyr Asp Arg Lys Thr Leu Asn Asn Asp Ile Met Leu Ile Lys Leu
                85                  90                  95
Ser Ser Arg Ala Val Ile Asn Ala Arg Val Ser Thr Ile Ser Leu Pro
            100                 105                 110
Thr Ala Pro Pro Ala Thr Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly
        115                 120                 125
Asn Thr Ile Thr Asn Ser Thr Ala Ser Ser Gly Ala Asp Tyr Pro Asp
    130                 135                 140
Glu Leu Gln Cys Leu Asp Ala Pro Val Leu Ser Gln Ala Lys Cys Glu
145                 150                 155                 160
Ala Ser Tyr Pro Gly Lys Ile Thr Ser Asn Met Phe Cys Val Gly Phe
                165                 170                 175
Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Val
            180                 185                 190
Val Cys Asn Gly Gln Leu Gln Gly Val Val Ser Trp Gly Asp Gly Cys
        195                 200                 205
Ala Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys Val Tyr Asn Tyr Val
    210                 215                 220
Lys Trp Ile Lys Asn Thr Ile Ala Ala Asn Ser
225                 230                 235
```

<210> 3
<211> 235
<212> PRT
<213> artificial sequence
<220>
<223> trypsin variant 2

<400> 3

```
Ile Val Gly Gly Tyr Asn Cys Glu Glu Asn Ser Val Pro Tyr Gln Val
1               5                   10                  15
Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Glu
            20                  25                  30
Gln Trp Val Val Ser Ala Gly His Cys Tyr Asn Gly Arg Asp Leu Glu
        35                  40                  45
Lys Ser Arg Ile Gln Val Arg Leu Gly Glu His Asn Ile Glu Val Leu
    50                  55                  60
Glu Gly Asn Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro
65                  70                  75                  80
Gln Tyr Asp Arg Lys Thr Leu Asn Asn Asp Ile Met Leu Ile Lys Leu
                85                  90                  95
Ser Ser Arg Ala Val Ile Asn Ala Arg Val Ser Thr Ile Ser Leu Pro
            100                 105                 110
Thr Ala Pro Pro Ala Thr Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly
        115                 120                 125
Asn Val Arg Gly Thr Trp Thr Ala Ser Ser Gly Ala Asp Tyr Pro Asp
    130                 135                 140
Glu Leu Gln Cys Leu Asp Ala Pro Val Leu Ser Gln Ala Lys Cys Glu
145                 150                 155                 160
Ala Ser Tyr Pro Gly Lys Ile Thr Ser Asn Met Phe Cys Val Gly Phe
                165                 170                 175
Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Val
            180                 185                 190
Val Cys Asn Gly Gln Leu Gln Gly Val Val Ser Trp Gly Asp Gly Cys
        195                 200                 205
Ala Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys Val Tyr Asn Tyr Val
    210                 215                 220
Lys Trp Ile Lys Asn Thr Ile Ala Ala Asn Ser
225                 230                 235
```

<210> 4
<211> 235
<212> PRT
<213> artificial sequence
<220>
<223> trypsin variant 3

<400> 4
Ile Val Gly Gly Tyr Asn Cys Glu Glu Asn Ser Val Pro Tyr Gln Val
1 5 10 15
Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Glu
20 25 30
Gln Trp Val Val Ser Ala Gly His Cys Tyr Ala Ala Thr Asn Gly Asp
35 40 45
Lys Ser Arg Ile Gln Val Arg Leu Gly Glu His Asn Ile Glu Val Leu
50 55 60
Glu Gly Asn Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro
65 70 75 80
Gln Tyr Asp Arg Lys Thr Leu Asn Asn Asp Ile Met Leu Ile Lys Leu
85 90 95
Ser Ser Arg Ala Val Ile Asn Ala Arg Val Ser Thr Ile Ser Leu Pro
100 105 110
Thr Ala Pro Pro Ala Thr Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly
115 120 125
Asn Arg Lys Asp Phe Trp Thr Ala Ser Ser Gly Ala Asp Tyr Pro Asp
130 135 140
Glu Leu Gln Cys Leu Asp Ala Pro Val Leu Ser Gln Ala Lys Cys Glu
145 150 155 160
Ala Ser Tyr Pro Gly Lys Ile Thr Ser Asn Met Phe Cys Val Gly Phe
165 170 175
Leu Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Val
180 185 190
Val Cys Asn Gly Gln Leu Gln Gly Val Val Ser Trp Gly Asp Gly Cys
195 200 205
Ala Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys Val Tyr Asn Tyr Val
210 215 220
Lys Trp Ile Lys Asn Thr Ile Ala Ala Asn Ser
225 230 235

<210> 5
<211> 259
<212> PRT
<213> Homo sapiens

<400> 5
Ile Val Glu Gly Ser Asp Ala Glu Ile Gly Met Ser Pro Trp Gln Val
1 5 10 15
Met Leu Phe Arg Lys Ser Pro Gln Glu Leu Leu Cys Gly Ala Ser Leu
20 25 30
Ile Ser Asp Arg Trp Val Leu Thr Ala Ala His Cys Leu Leu Tyr Pro
35 40 45
Pro Trp Asp Lys Asn Phe Thr Glu Asn Asp Leu Leu Val Arg Ile Gly
50 55 60
Lys His Ser Arg Thr Arg Tyr Glu Arg Asn Ile Glu Lys Ile Ser Met
65 70 75 80
Leu Glu Lys Ile Tyr Ile His Pro Arg Tyr Asn Trp Arg Glu Asn Leu
85 90 95
Asp Arg Asp Ile Ala Leu Met Lys Leu Lys Lys Pro Val Ala Phe Ser
100 105 110
Asp Tyr Ile His Pro Val Cys Leu Pro Asp Arg Glu Thr Ala Ala Ser
115 120 125

```
Leu Leu Gln Ala Gly Tyr Lys Gly Arg Val Thr Gly Trp Gly Asn Leu
    130                 135                 140
Lys Glu Thr Trp Thr Ala Asn Val Gly Lys Gly Gln Pro Ser Val Leu
145                 150                 155                 160
Gln Val Val Asn Leu Pro Ile Val Glu Arg Pro Val Cys Lys Asp Ser
                165                 170                 175
Thr Arg Ile Arg Ile Thr Asp Asn Met Phe Cys Ala Gly Tyr Lys Pro
            180                 185                 190
Asp Glu Gly Lys Arg Gly Asp Ala Cys Glu Gly Asp Ser Gly Gly Pro
        195                 200                 205
Phe Val Met Lys Ser Pro Phe Asn Asn Arg Trp Tyr Gln Met Gly Ile
    210                 215                 220
Val Ser Trp Gly Glu Gly Cys Asp Arg Asp Gly Lys Tyr Gly Phe Tyr
225                 230                 235                 240
Thr His Val Phe Arg Leu Lys Lys Trp Ile Gln Lys Val Ile Asp Gln
                245                 250                 255
Phe Gly Glu
```

<210> 6
<211> 235
<212> PRT
<213> Homo sapiens

```
<400> 6
Ile Val Gly Gly Ser Asn Ala Lys Glu Gly Ala Trp Pro Trp Val Val
1               5                   10                  15
Gly Leu Tyr Tyr Gly Gly Arg Leu Leu Cys Gly Ala Ser Leu Val Ser
            20                  25                  30
Ser Asp Trp Leu Val Ser Ala Ala His Cys Val Tyr Gly Arg Asn Leu
        35                  40                  45
Glu Pro Ser Lys Trp Thr Ala Ile Leu Gly Leu His Met Lys Ser Asn
    50                  55                  60
Leu Thr Ser Pro Gln Thr Val Pro Arg Leu Ile Asp Glu Ile Val Ile
65                  70                  75                  80
Asn Pro His Tyr Asn Arg Arg Arg Lys Asp Asn Asp Ile Ala Met Met
                85                  90                  95
His Leu Glu Phe Lys Val Asn Tyr Thr Asp Tyr Ile Gln Pro Ile Cys
            100                 105                 110
Leu Pro Glu Glu Asn Gln Val Phe Pro Pro Gly Arg Asn Cys Ser Ile
        115                 120                 125
Ala Gly Trp Gly Thr Val Val Tyr Gln Gly Thr Thr Ala Asn Ile Leu
    130                 135                 140
Gln Glu Ala Asp Val Pro Leu Leu Ser Asn Glu Arg Cys Gln Gln Gln
145                 150                 155                 160
Met Pro Glu Tyr Asn Ile Thr Glu Asn Met Ile Cys Ala Gly Tyr Glu
                165                 170                 175
Glu Gly Gly Ile Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Met
            180                 185                 190
Cys Gln Glu Asn Asn Arg Trp Phe Leu Ala Gly Val Thr Ser Phe Gly
        195                 200                 205
Tyr Lys Cys Ala Leu Pro Asn Arg Pro Gly Val Tyr Ala Arg Val Ser
    210                 215                 220
Arg Phe Thr Glu Trp Ile Gln Ser Phe Leu His
225                 230                 235
```

<210> 7
<211> 275
<212> PRT
<213> Bacillus subtilis

<400> 7
Ile Ala His Glu Tyr Ala Gln Ser Val Pro Tyr Gly Ile Ser Gln Ile
1 5 10 15
Lys Ala Pro Ala Leu His Ser Gln Gly Tyr Thr Gly Ser Asn Val Lys
20 25 30
Val Ala Val Ile Asp Ser Gly Ile Asp Ser Ser His Pro Asp Leu Asn
35 40 45
Val Arg Gly Gly Ala Ser Phe Val Pro Ser Glu Thr Asn Pro Tyr Gln
50 55 60
Asp Gly Ser Ser His Gly Thr His Val Ala Gly Thr Ile Ala Ala Leu
65 70 75 80
Asn Asn Ser Ile Gly Val Leu Gly Val Ser Pro Ser Ala Ser Leu Tyr
85 90 95
Ala Val Lys Val Leu Asp Ser Thr Gly Ser Gly Gln Tyr Ser Trp Ile
100 105 110
Ile Asn Gly Ile Glu Trp Ala Ile Ser Asn Asn Met Asp Val Ile Asn
115 120 125
Met Ser Leu Gly Gly Pro Thr Gly Ser Thr Ala Leu Lys Thr Val Val
130 135 140
Asp Lys Ala Val Ser Ser Gly Ile Val Val Ala Ala Ala Ala Gly Asn
145 150 155 160
Glu Gly Ser Ser Gly Ser Thr Ser Thr Val Gly Tyr Pro Ala Lys Tyr
165 170 175
Pro Ser Thr Ile Ala Val Gly Ala Val Asn Ser Ser Asn Gln Arg Ala
180 185 190
Ser Phe Ser Ser Ala Gly Ser Glu Leu Asp Val Met Ala Pro Gly Val
195 200 205
Ser Ile Gln Ser Thr Leu Pro Gly Gly Thr Tyr Gly Ala Tyr Asn Gly
210 215 220
Thr Ser Met Ala Thr Pro His Val Ala Gly Ala Ala Ala Leu Ile Leu
225 230 235 240
Ser Lys His Pro Thr Trp Thr Asn Ala Gln Val Arg Asp Arg Leu Glu
245 250 255
Ser Thr Ala Thr Tyr Leu Gly Asn Ser Phe Tyr Tyr Gly Lys Gly Leu
260 265 270
Ile Asn Val
275

<210> 8
<211> 320
<212> PRT
<213> Murinae gen. sp.

<400> 8
Val Ala Lys Arg Arg Ala Lys Arg Asp Val Tyr Gln Glu Pro Thr Asp
1 5 10 15
Pro Lys Phe Pro Gln Gln Trp Tyr Leu Ser Gly Val Thr Gln Arg Asp
20 25 30
Leu Asn Val Lys Glu Ala Trp Ala Gln Gly Phe Thr Gly His Gly Ile
35 40 45
Val Val Ser Ile Leu Asp Asp Gly Ile Glu Lys Asn His Pro Asp Leu
50 55 60
Ala Gly Asn Tyr Asp Pro Gly Ala Ser Phe Asp Val Asn Asp Gln Asp
65 70 75 80
Pro Asp Pro Gln Pro Arg Tyr Thr Gln Met Asn Asp Asn Arg His Gly
85 90 95
Thr Arg Cys Ala Gly Glu Val Ala Ala Val Ala Asn Asn Gly Val Cys
100 105 110
Gly Val Gly Val Ala Tyr Asn Ala Arg Ile Gly Gly Val Arg Met Leu
115 120 125
Asp Gly Glu Val Thr Asp Ala Val Glu Ala Arg Ser Leu Gly Leu Asn
130 135 140

```
Pro Asn His Ile His Ile Tyr Ser Ala Ser Trp Gly Pro Glu Asp Asp
145                 150                 155                 160
Gly Lys Thr Val Asp Gly Pro Ala Arg Leu Ala Glu Glu Ala Phe Phe
                165                 170                 175
Arg Gly Val Ser Gln Gly Arg Gly Gly Leu Gly Ser Ile Phe Val Trp
            180                 185                 190
Ala Ser Gly Asn Gly Gly Arg Glu His Asp Ser Cys Asn Cys Asp Gly
        195                 200                 205
Tyr Thr Asn Ser Ile Tyr Thr Leu Ser Ile Ser Ser Ala Thr Gln Phe
    210                 215                 220
Gly Asn Val Pro Trp Tyr Ser Glu Ala Cys Ser Ser Thr Leu Ala Thr
225                 230                 235                 240
Thr Tyr Ser Ser Gly Asn Gln Asn Glu Lys Gln Ile Val Thr Thr Asp
                245                 250                 255
Leu Arg Gln Lys Cys Thr Glu Ser His Thr Gly Thr Ser Ala Ser Ala
            260                 265                 270
Pro Leu Ala Ala Gly Ile Ile Ala Leu Thr Leu Glu Ala Asn Lys Asn
        275                 280                 285
Leu Thr Trp Arg Asp Met Gln His Leu Val Val Gln Thr Ser Lys Pro
    290                 295                 300
Ala His Leu Asn Ala Asp Asp Trp Ala Thr Asn Gly Val Gly Arg Lys
305                 310                 315                 320
```

<210> 9
<211> 330
<212> PRT
<213> Homo sapiens

<400> 9

```
Glu Lys Glu Arg Ser Lys Arg Ser Ala Leu Arg Asp Ser Ala Leu Asn
1               5                   10                  15
Leu Phe Asn Asp Pro Met Trp Asn Gln Gln Trp Tyr Leu Gln Asp Thr
            20                  25                  30
Arg Met Thr Ala Ala Leu Pro Lys Leu Asp Leu His Val Ile Pro Val
        35                  40                  45
Trp Gln Lys Gly Ile Thr Gly Lys Gly Val Val Ile Thr Val Leu Asp
    50                  55                  60
Asp Gly Leu Glu Trp Asn His Thr Asp Ile Tyr Ala Asn Tyr Asp Pro
65                  70                  75                  80
Glu Ala Ser Tyr Asp Phe Asn Asp Asn Asp His Asp Pro Phe Pro Arg
                85                  90                  95
Tyr Asp Pro Thr Asn Glu Asn Lys His Gly Thr Arg Cys Ala Gly Glu
            100                 105                 110
Ile Ala Met Gln Ala Asn Asn His Lys Cys Gly Val Gly Val Ala Tyr
        115                 120                 125
Asn Ser Lys Val Gly Gly Ile Arg Met Leu Asp Gly Ile Val Thr Asp
    130                 135                 140
Ala Ile Glu Ala Ser Ser Ile Gly Phe Asn Pro Gly His Val Asp Ile
145                 150                 155                 160
Tyr Ser Ala Ser Trp Gly Pro Asn Asp Asp Gly Lys Thr Val Glu Gly
                165                 170                 175
Pro Gly Arg Leu Ala Gln Lys Ala Phe Glu Tyr Gly Val Lys Gln Gly
            180                 185                 190
Arg Gln Gly Lys Gly Ser Ile Phe Val Trp Ala Ser Gly Asn Gly Gly
        195                 200                 205
Arg Gln Gly Asp Asn Cys Asp Cys Asp Gly Tyr Thr Asp Ser Ile Tyr
    210                 215                 220
Thr Ile Ser Ile Ser Ser Ala Ser Gln Gln Gly Leu Ser Pro Trp Tyr
225                 230                 235                 240
Ala Glu Lys Cys Ser Ser Thr Leu Ala Thr Ser Tyr Ser Ser Gly Asp
                245                 250                 255
Tyr Thr Asp Gln Arg Ile Thr Ser Ala Asp Leu His Asn Asp Cys Thr
```

```
            260                 265                 270
Glu Thr His Thr Gly Thr Ser Ala Ser Ala Pro Leu Ala Ala Gly Ile
        275                 280                 285
Phe Ala Leu Ala Leu Glu Ala Asn Pro Asn Leu Thr Trp Arg Asp Met
    290                 295                 300
Gln His Leu Val Val Trp Thr Ser Glu Tyr Asp Pro Leu Ala Asn Asn
305                 310                 315                 320
Pro Gly Trp Lys Lys Asn Gly Ala Gly Leu
                325                 330
```

<210> 10
<211> 297
<212> PRT
<213> Homo sapiens

<400> 10

```
Asn Thr His Pro Cys Gln Ser Asp Met Asn Ile Glu Gly Ala Trp Lys
1               5                   10                  15
Arg Gly Tyr Thr Gly Lys Asn Ile Val Val Thr Ile Leu Asp Asp Gly
            20                  25                  30
Ile Glu Arg Thr His Pro Asp Leu Met Gln Asn Tyr Asp Ala Leu Ala
        35                  40                  45
Ser Cys Asp Val Asn Gly Asn Asp Leu Asp Pro Met Pro Arg Tyr Asp
    50                  55                  60
Ala Ser Asn Glu Asn Lys His Gly Thr Arg Cys Ala Gly Glu Val Ala
65                  70                  75                  80
Ala Ala Ala Asn Asn Ser His Cys Thr Val Gly Ile Ala Phe Asn Ala
                85                  90                  95
Lys Ile Gly Gly Val Arg Met Leu Asp Gly Asp Val Thr Asp Met Val
            100                 105                 110
Glu Ala Lys Ser Val Ser Phe Asn Pro Gln His Val His Ile Tyr Ser
        115                 120                 125
Ala Ser Trp Gly Pro Asp Asp Asp Gly Lys Thr Val Asp Gly Pro Ala
    130                 135                 140
Pro Leu Thr Arg Gln Ala Phe Glu Asn Gly Val Arg Met Gly Arg Arg
145                 150                 155                 160
Gly Leu Gly Ser Val Phe Val Trp Ala Ser Gly Asn Gly Gly Arg Ser
                165                 170                 175
Lys Asp His Cys Ser Cys Asp Gly Tyr Thr Asn Ser Ile Tyr Thr Ile
            180                 185                 190
Ser Ile Ser Ser Thr Ala Glu Ser Gly Lys Lys Pro Trp Tyr Leu Glu
        195                 200                 205
Glu Cys Ser Ser Thr Leu Ala Thr Thr Tyr Ser Ser Gly Glu Ser Tyr
    210                 215                 220
Asp Lys Lys Ile Ile Thr Thr Asp Leu Arg Gln Arg Cys Thr Asp Asn
225                 230                 235                 240
His Thr Gly Thr Ser Ala Ser Ala Pro Met Ala Ala Gly Ile Ile Ala
                245                 250                 255
Leu Ala Leu Glu Ala Asn Pro Phe Leu Thr Trp Arg Asp Val Gln His
            260                 265                 270
Val Ile Val Arg Thr Ser Arg Ala Gly His Leu Asn Ala Asn Asp Trp
        275                 280                 285
Lys Thr Asn Ala Ala Gly Phe Lys Val
    290                 295
```

<210> 11
<211> 328
<212> PRT
<213> Homo sapiens

<400> 11

```
Thr Leu Val Asp Glu Gln Pro Leu Glu Asn Tyr Leu Asp Met Glu Tyr
1               5                  10                  15
Phe Gly Thr Ile Gly Ile Gly Thr Pro Ala Gln Asp Phe Thr Val Val
            20                  25                  30
Phe Asp Thr Gly Ser Ser Asn Leu Trp Val Pro Ser Val Tyr Cys Ser
        35                  40                  45
Ser Leu Ala Cys Thr Asn His Asn Arg Phe Asn Pro Glu Asp Ser Ser
    50                  55                  60
Thr Tyr Gln Ser Thr Ser Glu Thr Val Ser Ile Thr Tyr Gly Thr Gly
65                  70                  75                  80
Ser Met Thr Gly Ile Leu Gly Tyr Asp Thr Val Gln Val Gly Gly Ile
                85                  90                  95
Ser Asp Thr Asn Gln Ile Phe Gly Leu Ser Glu Thr Glu Pro Gly Ser
            100                 105                 110
Phe Leu Tyr Tyr Ala Pro Phe Asp Gly Ile Leu Gly Leu Ala Tyr Pro
        115                 120                 125
Ser Ile Ser Ser Ser Gly Ala Thr Pro Val Phe Asp Asn Ile Trp Asn
    130                 135                 140
Gln Gly Leu Val Ser Gln Asp Leu Phe Ser Val Tyr Leu Ser Ala Asp
145                 150                 155                 160
Asp Lys Ser Gly Ser Val Val Ile Phe Gly Gly Ile Asp Ser Ser Tyr
                165                 170                 175
Tyr Thr Gly Ser Leu Asn Trp Val Pro Val Thr Val Glu Gly Tyr Trp
            180                 185                 190
Gln Ile Thr Val Asp Ser Ile Thr Met Asn Gly Glu Thr Ile Ala Cys
        195                 200                 205
Ala Glu Gly Cys Gln Ala Ile Val Asp Thr Gly Thr Ser Leu Leu Thr
    210                 215                 220
Gly Pro Thr Ser Pro Ile Ala Asn Ile Gln Ser Asp Ile Gly Ala Ser
225                 230                 235                 240
Glu Asn Ser Asp Gly Asp Met Val Val Ser Cys Ser Ala Ile Ser Ser
                245                 250                 255
Leu Pro Asp Ile Val Phe Thr Ile Asn Gly Val Gln Tyr Pro Val Pro
            260                 265                 270
Pro Ser Ala Tyr Ile Leu Gln Ser Glu Gly Ser Cys Ile Ser Gly Phe
        275                 280                 285
Gln Gly Met Asn Val Pro Thr Glu Ser Gly Glu Leu Trp Ile Leu Gly
    290                 295                 300
Asp Val Phe Ile Arg Gln Tyr Phe Thr Val Phe Asp Arg Ala Asn Asn
305                 310                 315                 320
Gln Val Gly Leu Ala Pro Val Ala
                325
```

<210> 12
<211> 358
<212> PRT
<213> Homo sapiens

<400> 12

```
Glu Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val
1               5                  10                  15
Glu Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp
            20                  25                  30
Thr Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu
        35                  40                  45
His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg
    50                  55                  60
Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu
65                  70                  75                  80
Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg
                85                  90                  95
Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly
```

```
          100                 105                 110
Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg
        115                 120                 125
Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr
    130                 135                 140
His Val Pro Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro
145                 150                 155                 160
Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile
                165                 170                 175
Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro
            180                 185                 190
Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile
        195                 200                 205
Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys
    210                 215                 220
Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val
225                 230                 235                 240
Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys
                245                 250                 255
Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala
            260                 265                 270
Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met
        275                 280                 285
Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln
    290                 295                 300
Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr
305                 310                 315                 320
Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val
                325                 330                 335
Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile
            340                 345                 350
Gly Phe Ala Val Ser Ala
        355
```

<210> 13
<211> 351
<212> PRT
<213> Homo sapiens

<400> 13

```
Pro Ala Val Thr Glu Gly Pro Ile Pro Glu Val Leu Lys Asn Tyr Met
1               5                   10                  15
Asp Ala Gln Tyr Tyr Gly Glu Ile Gly Ile Gly Thr Pro Pro Gln Cys
            20                  25                  30
Phe Thr Val Val Phe Asp Thr Gly Ser Ser Asn Leu Trp Val Pro Ser
        35                  40                  45
Ile His Cys Lys Leu Leu Asp Ile Ala Cys Trp Ile His His Lys Tyr
    50                  55                  60
Asn Ser Asp Lys Ser Ser Thr Tyr Val Lys Asn Gly Thr Ser Phe Asp
65                  70                  75                  80
Ile His Tyr Gly Ser Gly Ser Leu Ser Gly Tyr Leu Ser Gln Asp Thr
                85                  90                  95
Val Ser Val Pro Cys Gln Ser Ala Ser Ser Ala Ser Ala Leu Gly Gly
            100                 105                 110
Val Lys Val Glu Arg Gln Val Phe Gly Glu Ala Thr Lys Gln Pro Gly
        115                 120                 125
Ile Thr Phe Ile Ala Ala Lys Phe Asp Gly Ile Leu Gly Met Ala Tyr
    130                 135                 140
Pro Arg Ile Ser Val Asn Asn Val Leu Pro Val Phe Asp Asn Leu Met
145                 150                 155                 160
Gln Gln Lys Leu Val Asp Gln Asn Ile Phe Ser Phe Tyr Leu Ser Arg
                165                 170                 175
```

```
Asp Pro Asp Ala Gln Pro Gly Gly Glu Leu Met Leu Gly Gly Thr Asp
            180                 185                 190
Ser Lys Tyr Tyr Lys Gly Ser Leu Ser Tyr Leu Asn Val Thr Arg Lys
        195                 200                 205
Ala Tyr Trp Gln Val His Leu Asp Gln Val Glu Val Ala Ser Gly Leu
    210                 215                 220
Thr Leu Cys Lys Glu Gly Cys Glu Ala Ile Val Asp Thr Gly Thr Ser
225                 230                 235                 240
Leu Met Val Gly Pro Val Asp Glu Val Arg Glu Leu Gln Lys Ala Ile
                245                 250                 255
Gly Ala Val Pro Leu Ile Gln Gly Glu Tyr Met Ile Pro Cys Glu Lys
            260                 265                 270
Val Ser Thr Leu Pro Ala Ile Thr Leu Lys Leu Gly Gly Lys Gly Tyr
        275                 280                 285
Lys Leu Ser Pro Glu Asp Tyr Thr Leu Lys Val Ser Gln Ala Gly Lys
    290                 295                 300
Thr Leu Cys Leu Ser Gly Phe Met Gly Met Asp Ile Pro Pro Pro Ser
305                 310                 315                 320
Gly Pro Leu Trp Ile Leu Gly Asp Val Phe Ile Gly Arg Tyr Tyr Thr
                325                 330                 335
Val Phe Asp Arg Asp Asn Asn Arg Val Gly Phe Ala Glu Ala Ala
            340                 345                 350
```

<210> 14
<211> 305
<212> PRT
<213> Homo sapiens

<400> 14

```
Met Leu Glu Ala Asp Asp Gln Gly Cys Ile Glu Glu Gln Gly Val Glu
1               5                   10                  15
Asp Ser Ala Asn Glu Asp Ser Val Asp Ala Lys Pro Asp Arg Ser Ser
            20                  25                  30
Phe Val Pro Ser Leu Phe Ser Lys Lys Lys Lys Asn Val Thr Met Arg
        35                  40                  45
Ser Ile Lys Thr Thr Arg Asp Arg Val Pro Thr Tyr Gln Tyr Asn Met
    50                  55                  60
Asn Phe Glu Lys Leu Gly Lys Cys Ile Ile Ile Asn Asn Lys Asn Phe
65                  70                  75                  80
Asp Lys Val Thr Gly Met Gly Val Arg Asn Gly Thr Asp Lys Asp Ala
                85                  90                  95
Glu Ala Leu Phe Lys Cys Phe Arg Ser Leu Gly Phe Asp Val Ile Val
            100                 105                 110
Tyr Asn Asp Cys Ser Cys Ala Lys Met Gln Asp Leu Leu Lys Lys Ala
        115                 120                 125
Ser Glu Glu Asp His Thr Asn Ala Ala Cys Phe Ala Cys Ile Leu Leu
    130                 135                 140
Ser His Gly Glu Glu Asn Val Ile Tyr Gly Lys Asp Gly Val Thr Pro
145                 150                 155                 160
Ile Lys Asp Leu Thr Ala His Phe Arg Gly Asp Arg Ser Lys Thr Leu
                165                 170                 175
Leu Glu Lys Pro Lys Leu Phe Phe Ile Gln Ala Cys Arg Gly Thr Glu
            180                 185                 190
Leu Asp Asp Gly Ile Gln Ala Asp Ser Gly Pro Ile Asn Asp Thr Asp
        195                 200                 205
Ala Asn Pro Arg Tyr Lys Ile Pro Val Glu Ala Asp Phe Leu Phe Ala
    210                 215                 220
Tyr Ser Thr Val Pro Gly Tyr Tyr Ser Trp Arg Ser Pro Gly Arg Gly
225                 230                 235                 240
Ser Trp Phe Val Gln Ala Leu Cys Ser Ile Leu Glu Glu His Gly Lys
                245                 250                 255
Asp Leu Glu Ile Met Gln Ile Leu Thr Arg Val Asn Asp Arg Val Ala
```

```
            260                 265                 270
Arg His Phe Glu Ser Gln Ser Asp Asp Pro His Phe His Glu Lys Lys
        275                 280                 285
Gln Ile Pro Cys Val Val Ser Met Leu Thr Lys Glu Leu Tyr Phe Ser
    290                 295                 300
Gln
305
```

<210> 15
<211> 262
<212> PRT
<213> Streptomyces sp. K15

<400> 15

```
Val Thr Lys Pro Thr Ile Ala Ala Val Gly Gly Tyr Ala Met Asn Asn
1               5                  10                  15
Gly Thr Gly Thr Thr Leu Tyr Thr Lys Ala Ala Asp Thr Arg Arg Ser
            20                  25                  30
Thr Gly Ser Thr Thr Lys Ile Met Thr Ala Lys Val Val Leu Ala Gln
        35                  40                  45
Ser Asn Leu Asn Leu Asp Ala Lys Val Thr Ile Gln Lys Ala Tyr Ser
    50                  55                  60
Asp Tyr Val Val Ala Asn Asn Ala Ser Gln Ala His Leu Ile Val Gly
65                  70                  75                  80
Asp Lys Val Thr Val Arg Gln Leu Leu Tyr Gly Leu Met Leu Pro Ser
                85                  90                  95
Gly Cys Asp Ala Ala Tyr Ala Leu Ala Asp Lys Tyr Gly Ser Gly Ser
            100                 105                 110
Thr Arg Ala Ala Arg Val Lys Ser Phe Ile Gly Lys Met Asn Thr Ala
        115                 120                 125
Ala Thr Asn Leu Gly Leu His Asn Thr His Phe Asp Ser Phe Asp Gly
    130                 135                 140
Ile Gly Asn Gly Ala Asn Tyr Ser Thr Pro Arg Asp Leu Thr Lys Ile
145                 150                 155                 160
Ala Ser Ser Ala Met Lys Asn Ser Thr Phe Arg Thr Val Val Lys Thr
                165                 170                 175
Lys Ala Tyr Thr Ala Lys Thr Val Thr Lys Thr Gly Ser Ile Arg Thr
            180                 185                 190
Met Asp Thr Trp Lys Asn Thr Asn Gly Leu Leu Ser Ser Tyr Ser Gly
        195                 200                 205
Ala Ile Gly Val Lys Thr Gly Ser Gly Pro Glu Ala Lys Tyr Cys Leu
    210                 215                 220
Val Phe Ala Ala Thr Arg Gly Gly Lys Thr Val Ile Gly Thr Val Leu
225                 230                 235                 240
Ala Ser Thr Ser Ile Pro Ala Arg Glu Ser Asp Ala Thr Lys Ile Met
                245                 250                 255
Asn Tyr Gly Phe Ala Leu
            260
```

<210> 16
<211> 256
<212> PRT
<213> Human cytomegalovirus

<400> 16

```
Met Thr Met Asp Glu Gln Gln Ser Gln Ala Val Ala Pro Val Tyr Val
1               5                  10                  15
Gly Gly Phe Leu Ala Arg Tyr Asp Gln Ser Pro Asp Glu Ala Glu Leu
            20                  25                  30
Leu Leu Pro Arg Asp Val Val Glu His Trp Leu His Ala Gln Gly Gln
        35                  40                  45
```

```
Gly Gln Pro Ser Leu Ser Val Ala Leu Pro Leu Asn Ile Asn His Asp
     50                  55                  60
Asp Thr Ala Val Val Gly His Val Ala Ala Met Gln Ser Val Arg Asp
65                   70                  75                  80
Gly Leu Phe Cys Leu Gly Cys Val Thr Ser Pro Arg Phe Leu Glu Ile
                 85                  90                  95
Val Arg Arg Ala Ser Glu Lys Ser Glu Leu Val Ser Arg Gly Pro Val
            100                 105                 110
Ser Pro Leu Gln Pro Asp Lys Val Val Glu Phe Leu Ser Gly Ser Tyr
        115                 120                 125
Ala Gly Leu Ser Leu Ser Ser Arg Arg Cys Asp Asp Val Glu Gln Ala
    130                 135                 140
Thr Ser Leu Ser Gly Ser Glu Thr Thr Pro Phe Lys His Val Ala Leu
145                 150                 155                 160
Cys Ser Val Gly Arg Arg Arg Gly Thr Leu Ala Val Tyr Gly Arg Asp
                165                 170                 175
Pro Glu Trp Val Thr Gln Arg Phe Pro Asp Leu Thr Ala Ala Asp Arg
            180                 185                 190
Asp Gly Leu Arg Ala Gln Trp Gln Arg Cys Gly Ser Thr Ala Val Asp
        195                 200                 205
Ala Ser Gly Asp Pro Phe Arg Ser Asp Ser Tyr Gly Leu Leu Gly Asn
    210                 215                 220
Ser Val Asp Ala Leu Tyr Ile Arg Glu Arg Leu Pro Lys Leu Arg Tyr
225                 230                 235                 240
Asp Lys Gln Leu Val Gly Val Thr Glu Arg Glu Ser Tyr Val Lys Ala
                245                 250                 255
```

<210> 17
<211> 248
<212> PRT
<213> Escherichia coli

<400> 17

```
Val Arg Ser Phe Ile Tyr Glu Pro Phe Gln Ile Pro Ser Gly Ser Met
1               5                  10                  15
Met Pro Thr Leu Leu Ile Gly Asp Phe Ile Leu Val Glu Lys Phe Ala
            20                  25                  30
Tyr Gly Ile Lys Asp Pro Ile Tyr Gln Lys Thr Leu Ile Glu Thr Gly
        35                  40                  45
His Pro Lys Arg Gly Asp Ile Val Val Phe Lys Tyr Pro Glu Asp Pro
    50                  55                  60
Lys Leu Asp Tyr Ile Lys Arg Ala Val Gly Leu Pro Gly Asp Lys Val
65                  70                  75                  80
Thr Tyr Asp Pro Val Ser Lys Glu Leu Thr Ile Gln Pro Gly Cys Ser
                85                  90                  95
Ser Gly Gln Ala Cys Glu Asn Ala Leu Pro Val Thr Tyr Ser Asn Val
            100                 105                 110
Glu Pro Ser Asp Phe Val Gln Thr Phe Ser Arg Arg Asn Gly Gly Glu
        115                 120                 125
Ala Thr Ser Gly Phe Phe Glu Val Pro Lys Asn Glu Thr Lys Glu Asn
    130                 135                 140
Gly Ile Arg Leu Ser Glu Arg Lys Glu Thr Leu Gly Asp Val Thr His
145                 150                 155                 160
Arg Ile Leu Thr Val Pro Ile Ala Gln Asp Gln Val Gly Met Tyr Tyr
                165                 170                 175
Gln Gln Pro Gly Gln Gln Leu Ala Thr Trp Ile Val Pro Pro Gly Gln
            180                 185                 190
Tyr Phe Met Met Gly Asp Asn Arg Asp Asn Ser Ala Asp Ser Arg Tyr
        195                 200                 205
Trp Gly Phe Val Pro Glu Ala Asn Leu Val Gly Arg Ala Thr Ala Ile
    210                 215                 220
Trp Met Ser Phe Asp Lys Gln Glu Gly Glu Trp Pro Thr Gly Leu Arg
```

225 230 235 240
Leu Ser Arg Ile Gly Gly Ile His
245

<210> 18
<211> 317
<212> PRT
<213> Serratia marcescens

<400> 18
Met Glu Gln Leu Arg Gly Leu Tyr Pro Pro Leu Ala Ala Tyr Asp Ser
1 5 10 15
Gly Trp Leu Asp Thr Gly Asp Gly His Arg Ile Tyr Trp Glu Leu Ser
20 25 30
Gly Asn Pro Asn Gly Lys Pro Ala Val Phe Ile His Gly Gly Pro Gly
35 40 45
Gly Gly Ile Ser Pro His His Arg Gln Leu Phe Asp Pro Glu Arg Tyr
50 55 60
Lys Val Leu Leu Phe Asp Gln Arg Gly Cys Gly Arg Ser Arg Pro His
65 70 75 80
Ala Ser Leu Asp Asn Asn Thr Thr Trp His Leu Val Ala Asp Ile Glu
85 90 95
Arg Leu Arg Glu Met Ala Gly Val Glu Gln Trp Leu Val Phe Gly Gly
100 105 110
Ser Trp Gly Ser Thr Leu Ala Leu Ala Tyr Ala Gln Thr His Pro Glu
115 120 125
Arg Val Ser Glu Met Val Leu Arg Gly Ile Phe Thr Leu Arg Lys Gln
130 135 140
Arg Leu His Trp Tyr Tyr Gln Asp Gly Ala Ser Arg Phe Phe Pro Glu
145 150 155 160
Lys Trp Glu Arg Val Leu Ser Ile Leu Ser Asp Asp Glu Arg Lys Asp
165 170 175
Val Ile Ala Ala Tyr Arg Gln Arg Leu Thr Ser Ala Asp Pro Gln Val
180 185 190
Gln Leu Glu Ala Ala Lys Leu Trp Ser Val Trp Glu Gly Glu Thr Val
195 200 205
Thr Leu Leu Pro Ser Arg Glu Ser Ala Ser Phe Gly Glu Asp Asp Phe
210 215 220
Ala Leu Ala Phe Ala Arg Ile Glu Asn His Tyr Phe Thr His Leu Gly
225 230 235 240
Phe Leu Glu Ser Asp Asp Gln Leu Leu Arg Asn Val Pro Leu Ile Arg
245 250 255
His Ile Pro Ala Val Ile Val His Gly Arg Tyr Asp Met Ala Cys Gln
260 265 270
Val Gln Asn Ala Trp Asp Leu Ala Lys Ala Trp Pro Glu Ala Glu Leu
275 280 285
His Ile Val Glu Gly Ala Gly His Ser Tyr Asp Glu Pro Gly Ile Leu
290 295 300
His Gln Leu Met Ile Ala Thr Asp Arg Phe Ala Gly Lys
305 310 315

<210> 19
<211> 229
<212> PRT
<213> Escherichia coli

<400> 19
Met Glu Leu Leu Leu Leu Ser Asn Ser Thr Leu Pro Gly Lys Ala Trp
1 5 10 15
Leu Glu His Ala Leu Pro Leu Ile Ala Asn Gln Leu Asn Gly Arg Arg
20 25 30

```
Ser Ala Val Phe Ile Pro Phe Ala Gly Val Thr Gln Thr Trp Asp Glu
        35                  40                  45
Tyr Thr Asp Lys Thr Ala Glu Val Leu Ala Pro Leu Gly Val Asn Val
    50                  55                  60
Thr Gly Ile His Arg Val Ala Asp Pro Leu Ala Ala Ile Glu Lys Ala
65                  70                  75                  80
Glu Ile Ile Ile Val Gly Gly Gly Asn Thr Phe Gln Leu Leu Lys Glu
                85                  90                  95
Ser Arg Glu Arg Gly Leu Leu Ala Pro Met Ala Asp Arg Val Lys Arg
            100                 105                 110
Gly Ala Leu Tyr Ile Gly Trp Ser Ala Gly Ala Asn Leu Ala Cys Pro
        115                 120                 125
Thr Ile Arg Thr Thr Asn Asp Met Pro Ile Val Asp Pro Asn Gly Phe
    130                 135                 140
Asp Ala Leu Asp Leu Phe Pro Leu Gln Ile Asn Pro His Phe Thr Asn
145                 150                 155                 160
Ala Leu Pro Glu Gly His Lys Gly Glu Thr Arg Glu Gln Arg Ile Arg
                165                 170                 175
Glu Leu Leu Val Val Ala Pro Glu Leu Thr Val Ile Gly Leu Pro Glu
            180                 185                 190
Gly Asn Trp Ile Gln Val Ser Asn Gly Gln Ala Val Leu Gly Gly Pro
        195                 200                 205
Asn Thr Thr Trp Val Phe Lys Ala Gly Glu Glu Ala Val Ala Leu Glu
    210                 215                 220
Ala Gly His Arg Phe
225
```

<210> 20
<211> 99
<212> PRT
<213> Human immunodeficiency virus

<400> 20

```
Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu Val Thr Val Lys Ile Gly
1               5                   10                  15
Gly Gln Leu Arg Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr Val
            20                  25                  30
Leu Glu Asp Ile Asn Leu Pro Gly Lys Trp Lys Pro Lys Met Ile Gly
        35                  40                  45
Gly Ile Gly Gly Phe Ile Lys Val Arg Gln Tyr Asp Gln Ile Leu Ile
    50                  55                  60
Glu Ile Cys Gly Lys Lys Ala Ile Gly Thr Val Leu Val Gly Pro Thr
65                  70                  75                  80
Pro Val Asn Ile Ile Gly Arg Asn Met Leu Thr Gln Ile Gly Cys Thr
                85                  90                  95
Leu Asn Phe
```

<210> 21
<211> 297
<212> PRT
<213> Escherichia coli

<400> 21

```
Ser Thr Glu Thr Leu Ser Phe Thr Pro Asp Asn Ile Asn Ala Asp Ile
1               5                   10                  15
Ser Leu Gly Thr Leu Ser Gly Lys Thr Lys Glu Arg Val Tyr Leu Ala
            20                  25                  30
Glu Glu Gly Gly Arg Lys Val Ser Gln Leu Asp Trp Lys Phe Asn Asn
        35                  40                  45
Ala Ala Ile Ile Lys Gly Ala Ile Asn Trp Asp Leu Met Pro Gln Ile
```

```
    50                  55                  60
Ser Ile Gly Ala Ala Gly Trp Thr Thr Leu Gly Ser Arg Gly Gly Asn
65                  70                  75                  80
Met Val Asp Gln Asp Trp Met Asp Ser Ser Asn Pro Gly Thr Trp Thr
                85                  90                  95
Asp Glu Ala Arg His Pro Asp Thr Gln Leu Asn Tyr Ala Asn Glu Phe
            100                 105                 110
Asp Leu Asn Ile Lys Gly Trp Leu Leu Asn Glu Pro Asn Tyr Arg Leu
        115                 120                 125
Gly Leu Met Ala Gly Tyr Gln Glu Ser Arg Tyr Ser Phe Thr Ala Arg
    130                 135                 140
Gly Gly Ser Tyr Ile Tyr Ser Ser Glu Glu Gly Phe Arg Asp Asp Ile
145                 150                 155                 160
Gly Ser Phe Pro Asn Gly Glu Arg Ala Ile Gly Tyr Lys Gln Arg Phe
                165                 170                 175
Lys Met Pro Tyr Ile Gly Leu Thr Gly Ser Tyr Arg Tyr Glu Asp Phe
            180                 185                 190
Glu Leu Gly Gly Thr Phe Lys Tyr Ser Gly Trp Val Glu Ser Ser Asp
        195                 200                 205
Asn Asp Glu His Tyr Asp Pro Lys Gly Arg Ile Thr Tyr Arg Ser Lys
    210                 215                 220
Val Lys Asp Gln Asn Tyr Tyr Ser Val Ala Val Asn Ala Gly Tyr Tyr
225                 230                 235                 240
Val Thr Pro Asn Ala Lys Val Tyr Val Glu Gly Ala Trp Asn Arg Val
                245                 250                 255
Thr Asn Lys Lys Gly Asn Thr Ser Leu Tyr Asp His Asn Asn Asn Thr
            260                 265                 270
Ser Asp Tyr Ser Lys Asn Gly Ala Gly Ile Glu Asn Tyr Asn Phe Ile
        275                 280                 285
Thr Thr Ala Gly Leu Lys Tyr Thr Phe
    290                 295
```

<210> 22
<211> 212
<212> PRT
<213> Carica papaya

<400> 22

```
Ile Pro Glu Tyr Val Asp Trp Arg Gln Lys Gly Ala Val Thr Pro Val
1               5                   10                  15
Lys Asn Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Ser Ala Val Val
            20                  25                  30
Thr Ile Glu Gly Ile Ile Lys Ile Arg Thr Gly Asn Leu Asn Gln Tyr
        35                  40                  45
Ser Glu Gln Glu Leu Leu Asp Cys Asp Arg Arg Ser Tyr Gly Cys Asn
    50                  55                  60
Gly Gly Tyr Pro Trp Ser Ala Leu Gln Leu Val Ala Gln Tyr Gly Ile
65                  70                  75                  80
His Tyr Arg Asn Thr Tyr Pro Tyr Glu Gly Val Gln Arg Tyr Cys Arg
                85                  90                  95
Ser Arg Glu Lys Gly Pro Tyr Ala Ala Lys Thr Asp Gly Val Arg Gln
            100                 105                 110
Val Gln Pro Tyr Asn Gln Gly Ala Leu Leu Tyr Ser Ile Ala Asn Gln
        115                 120                 125
Pro Val Ser Val Val Leu Gln Ala Ala Gly Lys Asp Phe Gln Leu Tyr
    130                 135                 140
Arg Gly Gly Ile Phe Val Gly Pro Cys Gly Asn Lys Val Asp His Ala
145                 150                 155                 160
Val Ala Ala Val Gly Tyr Gly Pro Asn Tyr Ile Leu Ile Lys Asn Ser
                165                 170                 175
Trp Gly Thr Gly Trp Gly Glu Asn Gly Tyr Ile Arg Ile Lys Arg Gly
            180                 185                 190
```

Thr Gly Asn Ser Tyr Gly Val Cys Gly Leu Tyr Thr Ser Ser Phe Tyr
195 200 205
Pro Val Lys Asn
210

<210> 23
<211> 699
<212> PRT
<213> Homo sapiens

<400> 23
Ala Gly Ile Ala Ala Lys Leu Ala Lys Asp Arg Glu Ala Ala Glu Gly
1 5 10 15
Leu Gly Ser His Glu Arg Ala Ile Lys Tyr Leu Asn Gln Asp Tyr Glu
20 25 30
Ala Leu Arg Asn Glu Cys Leu Glu Ala Gly Thr Leu Phe Gln Asp Pro
35 40 45
Ser Phe Pro Ala Ile Pro Ser Ala Leu Gly Phe Lys Glu Leu Gly Pro
50 55 60
Tyr Ser Ser Lys Thr Arg Gly Met Arg Trp Lys Arg Pro Thr Glu Ile
65 70 75 80
Cys Ala Asp Pro Gln Phe Ile Ile Gly Gly Ala Thr Arg Thr Asp Ile
85 90 95
Cys Gln Gly Ala Leu Gly Asp Cys Trp Leu Leu Ala Ala Ile Ala Ser
100 105 110
Leu Thr Leu Asn Glu Glu Ile Leu Ala Arg Val Val Pro Leu Asn Gln
115 120 125
Ser Phe Gln Glu Asn Tyr Ala Gly Ile Phe His Phe Gln Phe Trp Gln
130 135 140
Tyr Gly Glu Trp Val Glu Val Val Val Asp Asp Arg Leu Pro Thr Lys
145 150 155 160
Asp Gly Glu Leu Leu Phe Val His Ser Ala Glu Gly Ser Glu Phe Trp
165 170 175
Ser Ala Leu Leu Glu Lys Ala Tyr Ala Lys Ile Asn Gly Cys Tyr Glu
180 185 190
Ala Leu Ser Gly Gly Ala Thr Thr Glu Gly Phe Glu Asp Phe Thr Gly
195 200 205
Gly Ile Ala Glu Trp Tyr Glu Leu Lys Lys Pro Pro Pro Asn Leu Phe
210 215 220
Lys Ile Ile Gln Lys Ala Leu Gln Lys Gly Ser Leu Leu Gly Cys Ser
225 230 235 240
Ile Asp Ile Thr Ser Ala Ala Asp Ser Glu Ala Ile Thr Phe Gln Lys
245 250 255
Leu Val Lys Gly His Ala Tyr Ser Val Thr Gly Ala Glu Glu Val Glu
260 265 270
Ser Asn Gly Ser Leu Gln Lys Leu Ile Arg Ile Arg Asn Pro Trp Gly
275 280 285
Glu Val Glu Trp Thr Gly Arg Trp Asn Asp Asn Cys Pro Ser Trp Asn
290 295 300
Thr Ile Asp Pro Glu Glu Arg Glu Arg Leu Thr Arg Arg His Glu Asp
305 310 315 320
Gly Glu Phe Trp Met Ser Phe Ser Asp Phe Leu Arg His Tyr Ser Arg
325 330 335
Leu Glu Ile Cys Asn Leu Thr Pro Asp Thr Leu Thr Ser Asp Thr Tyr
340 345 350
Lys Lys Trp Lys Leu Thr Lys Met Asp Gly Asn Trp Arg Arg Gly Ser
355 360 365
Thr Ala Gly Gly Cys Arg Asn Tyr Pro Asn Thr Phe Trp Met Asn Pro
370 375 380
Gln Tyr Leu Ile Lys Leu Glu Glu Glu Asp Glu Asp Glu Glu Asp Gly
385 390 395 400
Glu Ser Gly Cys Thr Phe Leu Val Gly Leu Ile Gln Lys His Arg Arg

```
                405                 410                 415
Arg Gln Arg Lys Met Gly Glu Asp Met His Thr Ile Gly Phe Gly Ile
            420                 425                 430
Tyr Glu Val Pro Glu Glu Leu Ser Gly Gln Thr Asn Ile His Leu Ser
        435                 440                 445
Lys Asn Phe Phe Leu Thr Asn Arg Ala Arg Glu Arg Ser Asp Thr Phe
    450                 455                 460
Ile Asn Leu Arg Glu Val Leu Asn Arg Phe Lys Leu Pro Pro Gly Glu
465                 470                 475                 480
Tyr Ile Leu Val Pro Ser Thr Phe Glu Pro Asn Lys Asp Gly Asp Phe
                485                 490                 495
Cys Ile Arg Val Phe Ser Glu Lys Lys Ala Asp Tyr Gln Ala Val Asp
            500                 505                 510
Asp Glu Ile Glu Ala Asn Leu Glu Glu Phe Asp Ile Ser Glu Asp Asp
        515                 520                 525
Ile Asp Asp Gly Val Arg Arg Leu Phe Ala Gln Leu Ala Gly Glu Asp
    530                 535                 540
Ala Glu Ile Ser Ala Phe Glu Leu Gln Thr Ile Leu Arg Arg Val Leu
545                 550                 555                 560
Ala Lys Arg Gln Asp Ile Lys Ser Asp Gly Phe Ser Ile Glu Thr Cys
                565                 570                 575
Lys Ile Met Val Asp Met Leu Asp Ser Asp Gly Ser Gly Lys Leu Gly
            580                 585                 590
Leu Lys Glu Phe Tyr Ile Leu Trp Thr Lys Ile Gln Lys Tyr Gln Lys
        595                 600                 605
Ile Tyr Arg Glu Ile Asp Val Asp Arg Ser Gly Thr Met Asn Ser Tyr
    610                 615                 620
Glu Met Arg Lys Ala Leu Glu Glu Ala Gly Phe Lys Met Pro Cys Gln
625                 630                 635                 640
Leu His Gln Val Ile Val Ala Arg Phe Ala Asp Asp Gln Leu Ile Ile
                645                 650                 655
Asp Phe Asp Asn Phe Val Arg Cys Leu Val Arg Leu Glu Thr Leu Phe
            660                 665                 670
Lys Ile Phe Lys Gln Leu Asp Pro Glu Asn Thr Gly Thr Ile Glu Leu
        675                 680                 685
Asp Leu Ile Ser Trp Leu Cys Phe Ser Val Leu
    690                 695
```

<210> 24
<211> 221
<212> PRT
<213> Tobacco etch virus

<400> 24

```
Gly Glu Ser Leu Phe Lys Gly Pro Arg Asp Tyr Asn Pro Ile Ser Ser
1               5                   10                  15
Thr Ile Cys His Leu Thr Asn Glu Ser Asp Gly His Thr Thr Ser Leu
            20                  25                  30
Tyr Gly Ile Gly Phe Gly Pro Phe Ile Ile Thr Asn Lys His Leu Phe
        35                  40                  45
Arg Arg Asn Asn Gly Thr Leu Leu Val Gln Ser Leu His Gly Val Phe
    50                  55                  60
Lys Val Lys Asn Thr Thr Thr Leu Gln Gln His Leu Ile Asp Gly Arg
65                  70                  75                  80
Asp Met Ile Ile Ile Arg Met Pro Lys Asp Phe Pro Pro Phe Pro Gln
                85                  90                  95
Lys Leu Lys Phe Arg Glu Pro Gln Arg Glu Glu Arg Ile Cys Leu Val
            100                 105                 110
Thr Thr Asn Phe Gln Thr Lys Ser Met Ser Ser Met Val Ser Asp Thr
        115                 120                 125
Ser Cys Thr Phe Pro Ser Ser Asp Gly Ile Phe Trp Lys His Trp Ile
    130                 135                 140
```

```
Gln Thr Lys Asp Gly Gln Cys Gly Ser Pro Leu Val Ser Thr Arg Asp
145                 150                 155                 160
Gly Phe Ile Val Gly Ile His Ser Ala Ser Asn Phe Thr Asn Thr Asn
                165                 170                 175
Asn Tyr Phe Thr Ser Val Pro Lys Asn Phe Met Glu Leu Leu Thr Asn
            180                 185                 190
Gln Glu Ala Gln Gln Trp Val Ser Gly Trp Arg Leu Asn Ala Asp Ser
        195                 200                 205
Val Leu Trp Gly Gly His Lys Val Phe Met Asp Lys Pro
    210                 215                 220
```

<210> 25
<211> 371
<212> PRT
<213> Streptococcus pyogenes

<400> 25

```
Asp Gln Asn Phe Ala Arg Asn Glu Lys Glu Ala Lys Asp Ser Ala Ile
1               5                  10                  15
Thr Phe Ile Gln Lys Ser Ala Ala Ile Lys Ala Gly Ala Arg Ser Ala
            20                  25                  30
Glu Asp Ile Lys Leu Asp Lys Val Asn Leu Gly Gly Glu Leu Ser Gly
        35                  40                  45
Ser Asn Met Tyr Val Tyr Asn Ile Ser Thr Gly Gly Phe Val Ile Val
    50                  55                  60
Ser Gly Asp Lys Arg Ser Pro Glu Ile Leu Gly Tyr Ser Thr Ser Gly
65                  70                  75                  80
Ser Phe Asp Val Asn Gly Lys Glu Asn Ile Ala Ser Phe Met Glu Ser
                85                  90                  95
Tyr Val Glu Gln Ile Lys Glu Asn Lys Lys Leu Asp Ser Thr Tyr Ala
            100                 105                 110
Gly Thr Ala Glu Ile Lys Gln Pro Val Val Lys Ser Leu Leu Asp Ser
        115                 120                 125
Lys Gly Ile His Tyr Asn Gln Gly Asn Pro Tyr Asn Leu Leu Thr Pro
    130                 135                 140
Val Ile Glu Lys Val Lys Pro Gly Glu Gln Ser Phe Val Gly Gln His
145                 150                 155                 160
Ala Ala Thr Gly Ser Val Ala Thr Ala Thr Ala Gln Ile Met Lys Tyr
                165                 170                 175
His Asn Tyr Pro Asn Lys Gly Leu Lys Asp Tyr Thr Tyr Thr Leu Ser
            180                 185                 190
Ser Asn Asn Pro Tyr Phe Asn His Pro Lys Asn Leu Phe Ala Ala Ile
        195                 200                 205
Ser Thr Arg Gln Tyr Asn Trp Asn Asn Ile Leu Pro Thr Tyr Ser Gly
    210                 215                 220
Arg Glu Ser Asn Val Gln Lys Met Ala Ile Ser Glu Leu Met Ala Asp
225                 230                 235                 240
Val Gly Ile Ser Val Asp Met Asp Tyr Gly Pro Ser Ser Gly Ser Ala
                245                 250                 255
Gly Ser Ser Arg Val Gln Arg Ala Leu Lys Glu Asn Phe Gly Tyr Asn
            260                 265                 270
Gln Ser Val His Gln Ile Asn Arg Gly Asp Phe Ser Lys Gln Asp Trp
        275                 280                 285
Glu Ala Gln Ile Asp Lys Glu Leu Ser Gln Asn Gln Pro Val Tyr Tyr
    290                 295                 300
Gln Gly Val Gly Lys Val Gly Gly His Ala Phe Val Ile Asp Gly Ala
305                 310                 315                 320
Asp Gly Arg Asn Phe Tyr His Val Asn Trp Gly Trp Gly Gly Val Ser
                325                 330                 335
Asp Gly Phe Phe Arg Leu Asp Ala Leu Asn Pro Ser Ala Leu Gly Thr
            340                 345                 350
Gly Gly Gly Ala Gly Gly Phe Asn Gly Tyr Gln Ser Ala Val Val Gly
        355                 360                 365
```

Ile Lys Pro
370

<210> 26
<211> 353
<212> PRT
<213> Homo sapiens

<400> 26
Lys Lys His Thr Gly Tyr Val Gly Leu Lys Asn Gln Gly Ala Thr Cys
1 5 10 15
Tyr Met Asn Ser Leu Leu Gln Thr Leu Phe Phe Thr Asn Gln Leu Arg
20 25 30
Lys Ala Val Tyr Met Met Pro Thr Glu Gly Asp Asp Ser Ser Lys Ser
35 40 45
Val Pro Leu Ala Leu Gln Arg Val Phe Tyr Glu Leu Gln His Ser Asp
50 55 60
Lys Pro Val Gly Thr Lys Lys Leu Thr Lys Ser Phe Gly Trp Glu Thr
65 70 75 80
Leu Asp Ser Phe Met Gln His Asp Val Gln Glu Leu Cys Arg Val Leu
85 90 95
Leu Asp Asn Val Glu Asn Lys Met Lys Gly Thr Cys Val Glu Gly Thr
100 105 110
Ile Pro Lys Leu Phe Arg Gly Lys Met Val Ser Tyr Ile Gln Cys Lys
115 120 125
Glu Val Asp Tyr Arg Ser Asp Arg Arg Glu Asp Tyr Tyr Asp Ile Gln
130 135 140
Leu Ser Ile Lys Gly Lys Lys Asn Ile Phe Glu Ser Phe Val Asp Tyr
145 150 155 160
Val Ala Val Glu Gln Leu Asp Gly Asp Asn Lys Tyr Asp Ala Gly Glu
165 170 175
His Gly Leu Gln Glu Ala Glu Lys Gly Val Lys Phe Leu Thr Leu Pro
180 185 190
Pro Val Leu His Leu Gln Leu Met Arg Phe Met Tyr Asp Pro Gln Thr
195 200 205
Asp Gln Asn Ile Lys Ile Asn Asp Arg Phe Glu Phe Pro Glu Gln Leu
210 215 220
Pro Leu Asp Glu Phe Leu Gln Lys Thr Asp Pro Lys Asp Pro Ala Asn
225 230 235 240
Tyr Ile Leu His Ala Val Leu Val His Ser Gly Asp Asn His Gly Gly
245 250 255
His Tyr Val Val Tyr Leu Asn Pro Lys Gly Asp Gly Lys Trp Cys Lys
260 265 270
Phe Asp Asp Asp Val Val Ser Arg Cys Thr Lys Glu Glu Ala Ile Glu
275 280 285
His Asn Tyr Gly Gly His Asp Asp Asp Leu Ser Val Arg His Cys Thr
290 295 300
Asn Ala Tyr Met Leu Val Tyr Ile Arg Glu Ser Lys Leu Ser Glu Val
305 310 315 320
Leu Gln Ala Val Thr Asp His Asp Ile Pro Gln Gln Leu Val Glu Arg
325 330 335
Leu Gln Glu Glu Lys Arg Ile Glu Ala Gln Lys Arg Lys Glu Arg Gln
340 345 350
Glu

<210> 27
<211> 174
<212> PRT
<213> Staphylococcus aureus

<400> 27
Tyr Asn Glu Gln Tyr Val Asn Lys Leu Glu Asn Phe Lys Ile Arg Glu
1 5 10 15
Thr Gln Gly Asn Asn Gly Trp Cys Ala Gly Tyr Thr Met Ser Ala Leu
20 25 30
Leu Asn Ala Thr Tyr Asn Thr Asn Lys Tyr His Ala Glu Ala Val Met
35 40 45
Arg Phe Leu His Pro Asn Leu Gln Gly Gln Gln Phe Gln Phe Thr Gly
50 55 60
Leu Thr Pro Arg Glu Met Ile Tyr Phe Gly Gln Thr Gln Gly Arg Ser
65 70 75 80
Pro Gln Leu Leu Asn Arg Met Thr Thr Tyr Asn Glu Val Asp Asn Leu
85 90 95
Thr Lys Asn Asn Lys Gly Ile Ala Ile Leu Gly Ser Arg Val Glu Ser
100 105 110
Arg Asn Gly Met His Ala Gly His Ala Met Ala Val Val Gly Asn Ala
115 120 125
Lys Leu Asn Asn Gly Gln Glu Val Ile Ile Ile Trp Asn Pro Trp Asp
130 135 140
Asn Gly Phe Met Thr Gln Asp Ala Lys Asn Asn Val Ile Pro Val Ser
145 150 155 160
Asn Gly Asp His Tyr Gln Trp Tyr Ser Ser Ile Tyr Gly Tyr
165 170

<210> 28
<211> 221
<212> PRT
<213> Saccharomyces cerevisiae

<400> 28
Gly Ser Leu Val Pro Glu Leu Asn Glu Lys Asp Asp Asp Gln Val Gln
1 5 10 15
Lys Ala Leu Ala Ser Arg Glu Asn Thr Gln Leu Met Asn Arg Asp Asn
20 25 30
Ile Glu Ile Thr Val Arg Asp Phe Lys Thr Leu Ala Pro Arg Arg Trp
35 40 45
Leu Asn Asp Thr Ile Ile Glu Phe Phe Met Lys Tyr Ile Glu Lys Ser
50 55 60
Thr Pro Asn Thr Val Ala Phe Asn Ser Phe Phe Tyr Thr Asn Leu Ser
65 70 75 80
Glu Arg Gly Tyr Gln Gly Val Arg Arg Trp Met Lys Arg Lys Lys Thr
85 90 95
Gln Ile Asp Lys Leu Asp Lys Ile Phe Thr Pro Ile Asn Leu Asn Gln
100 105 110
Ser His Trp Ala Leu Gly Ile Ile Asp Leu Lys Lys Lys Thr Ile Gly
115 120 125
Tyr Val Asp Ser Leu Ser Asn Gly Pro Asn Ala Met Ser Phe Ala Ile
130 135 140
Leu Thr Asp Leu Gln Lys Tyr Val Met Glu Glu Ser Lys His Thr Ile
145 150 155 160
Gly Glu Asp Phe Asp Leu Ile His Leu Asp Cys Pro Gln Gln Pro Asn
165 170 175
Gly Tyr Asp Cys Gly Ile Tyr Val Cys Met Asn Thr Leu Tyr Gly Ser
180 185 190
Ala Asp Ala Pro Leu Asp Phe Asp Tyr Lys Asp Ala Ile Arg Met Arg
195 200 205
Arg Phe Ile Ala His Leu Ile Leu Thr Asp Ala Leu Lys
210 215 220

<210> 29
<211> 166

<212> PRT
<213> Pyrococcus horikoshii

<400> 29
Met Lys Val Leu Phe Leu Thr Ala Asn Glu Phe Glu Asp Val Glu Leu
1 5 10 15
Ile Tyr Pro Tyr His Arg Leu Lys Glu Glu Gly His Glu Val Tyr Ile
20 25 30
Ala Ser Phe Glu Arg Gly Thr Ile Thr Gly Lys His Gly Tyr Ser Val
35 40 45
Lys Val Asp Leu Thr Phe Asp Lys Val Asn Pro Glu Glu Phe Asp Ala
50 55 60
Leu Val Leu Pro Gly Gly Arg Ala Pro Glu Arg Val Arg Leu Asn Glu
65 70 75 80
Lys Ala Val Ser Ile Ala Arg Lys Met Phe Ser Glu Gly Lys Pro Val
85 90 95
Ala Ser Ile Cys His Gly Pro Gln Ile Leu Ile Ser Ala Gly Val Leu
100 105 110
Arg Gly Arg Lys Gly Thr Ser Tyr Pro Gly Ile Lys Asp Asp Met Ile
115 120 125
Asn Ala Gly Val Glu Trp Val Asp Ala Glu Val Val Val Asp Gly Asn
130 135 140
Trp Val Ser Ser Arg Val Pro Ala Asp Leu Tyr Ala Trp Met Arg Glu
145 150 155 160
Phe Val Lys Leu Leu Lys
165

<210> 30
<211> 316
<212> PRT
<213> Bacillus thermoproteolyticus

<400> 30
Ile Thr Gly Thr Ser Thr Val Gly Val Gly Arg Gly Val Leu Gly Asp
1 5 10 15
Gln Lys Asn Ile Asn Thr Thr Tyr Ser Thr Tyr Tyr Tyr Leu Gln Asp
20 25 30
Asn Thr Arg Gly Asp Gly Ile Phe Thr Tyr Asp Ala Lys Tyr Arg Thr
35 40 45
Thr Leu Pro Gly Ser Leu Trp Ala Asp Ala Asp Asn Gln Phe Phe Ala
50 55 60
Ser Tyr Asp Ala Pro Ala Val Asp Ala His Tyr Tyr Ala Gly Val Thr
65 70 75 80
Tyr Asp Tyr Tyr Lys Asn Val His Asn Arg Leu Ser Tyr Asp Gly Asn
85 90 95
Asn Ala Ala Ile Arg Ser Ser Val His Tyr Ser Gln Gly Tyr Asn Asn
100 105 110
Ala Phe Trp Asn Gly Ser Glu Met Val Tyr Gly Asp Gly Asp Gly Gln
115 120 125
Thr Phe Ile Pro Leu Ser Gly Gly Ile Asp Val Val Ala His Glu Leu
130 135 140
Thr His Ala Val Thr Asp Tyr Thr Ala Gly Leu Ile Tyr Gln Asn Glu
145 150 155 160
Ser Gly Ala Ile Asn Glu Ala Ile Ser Asp Ile Phe Gly Thr Leu Val
165 170 175
Glu Phe Tyr Ala Asn Lys Asn Pro Asp Trp Glu Ile Gly Glu Asp Val
180 185 190
Tyr Thr Pro Gly Ile Ser Gly Asp Ser Leu Arg Ser Met Ser Asp Pro
195 200 205
Ala Lys Tyr Gly Asp Pro Asp His Tyr Ser Lys Arg Tyr Thr Gly Thr
210 215 220
Gln Asp Asn Gly Gly Val His Ile Asn Ser Gly Ile Ile Asn Lys Ala

225 230 235 240
Ala Tyr Leu Ile Ser Gln Gly Gly Thr His Tyr Gly Val Ser Val Val
245 250 255
Gly Ile Gly Arg Asp Lys Leu Gly Lys Ile Phe Tyr Arg Ala Leu Thr
260 265 270
Gln Tyr Leu Thr Pro Thr Ser Asn Phe Ser Gln Leu Arg Ala Ala Ala
275 280 285
Val Gln Ser Ala Thr Asp Leu Tyr Gly Ser Thr Ser Gln Glu Val Ala
290 295 300
Ser Val Lys Gln Ala Phe Asp Ala Val Gly Val Lys
305 310 315

<210> 31
<211> 169
<212> PRT
<213> Homo sapiens

<400> 31
Val Leu Thr Glu Gly Asn Pro Arg Trp Glu Gln Thr His Leu Thr Tyr
1 5 10 15
Arg Ile Glu Asn Tyr Thr Pro Asp Leu Pro Arg Ala Asp Val Asp His
20 25 30
Ala Ile Glu Lys Ala Phe Gln Leu Trp Ser Asn Val Thr Pro Leu Thr
35 40 45
Phe Thr Lys Val Ser Glu Gly Gln Ala Asp Ile Met Ile Ser Phe Val
50 55 60
Arg Gly Asp His Arg Asp Asn Ser Pro Phe Asp Gly Pro Gly Gly Asn
65 70 75 80
Leu Ala His Ala Phe Gln Pro Gly Pro Gly Ile Gly Gly Asp Ala His
85 90 95
Phe Asp Glu Asp Glu Arg Trp Thr Asn Asn Phe Arg Glu Tyr Asn Leu
100 105 110
His Arg Val Ala Ala His Glu Leu Gly His Ser Leu Gly Leu Ser His
115 120 125
Ser Thr Asp Ile Gly Ala Leu Met Tyr Pro Ser Tyr Thr Phe Ser Gly
130 135 140
Asp Val Gln Leu Ala Gln Asp Asp Ile Asp Gly Ile Gln Ala Ile Tyr
145 150 155 160
Gly Arg Ser Gln Asn Pro Val Gln Pro
165

<210> 32
<211> 496
<212> PRT
<213> Homo sapiens

<400> 32
Gln Tyr Ser Pro Asn Thr Gln Gln Gly Arg Thr Ser Ile Val His Leu
1 5 10 15
Phe Glu Trp Arg Trp Val Asp Ile Ala Leu Glu Cys Glu Arg Tyr Leu
20 25 30
Ala Pro Lys Gly Phe Gly Gly Val Gln Val Ser Pro Pro Asn Glu Asn
35 40 45
Val Ala Ile Tyr Asn Pro Phe Arg Pro Trp Trp Glu Arg Tyr Gln Pro
50 55 60
Val Ser Tyr Lys Leu Cys Thr Arg Ser Gly Asn Glu Asp Glu Phe Arg
65 70 75 80
Asn Met Val Thr Arg Cys Asn Asn Val Gly Val Arg Ile Tyr Val Asp
85 90 95
Ala Val Ile Asn His Met Cys Gly Asn Ala Val Ser Ala Gly Thr Ser
100 105 110

```
Ser Thr Cys Gly Ser Tyr Phe Asn Pro Gly Ser Arg Asp Phe Pro Ala
        115                 120                 125
Val Pro Tyr Ser Gly Trp Asp Phe Asn Asp Gly Lys Cys Lys Thr Gly
    130                 135                 140
Ser Gly Asp Ile Glu Asn Tyr Asn Asp Ala Thr Gln Val Arg Asp Cys
145                 150                 155                 160
Arg Leu Thr Gly Leu Leu Asp Leu Ala Leu Glu Lys Asp Tyr Val Arg
                165                 170                 175
Ser Lys Ile Ala Glu Tyr Met Asn His Leu Ile Asp Ile Gly Val Ala
            180                 185                 190
Gly Phe Arg Leu Asp Ala Ser Lys His Met Trp Pro Gly Asp Ile Lys
        195                 200                 205
Ala Ile Leu Asp Lys Leu His Asn Leu Asn Ser Asn Trp Phe Pro Ala
    210                 215                 220
Gly Ser Lys Pro Phe Ile Tyr Gln Glu Val Ile Asp Leu Gly Gly Glu
225                 230                 235                 240
Pro Ile Lys Ser Ser Asp Tyr Phe Gly Asn Gly Arg Val Thr Glu Phe
                245                 250                 255
Lys Tyr Gly Ala Lys Leu Gly Thr Val Ile Arg Lys Trp Asn Gly Glu
            260                 265                 270
Lys Met Ser Tyr Leu Lys Asn Trp Gly Glu Gly Trp Gly Phe Val Pro
        275                 280                 285
Ser Asp Arg Ala Leu Val Phe Val Asp Asn His Asp Asn Gln Arg Gly
    290                 295                 300
His Gly Ala Gly Gly Ala Ser Ile Leu Thr Phe Trp Asp Ala Arg Leu
305                 310                 315                 320
Tyr Lys Met Ala Val Gly Phe Met Leu Ala His Pro Tyr Gly Phe Thr
                325                 330                 335
Arg Val Met Ser Ser Tyr Arg Trp Pro Arg Gln Phe Gln Asn Gly Asn
            340                 345                 350
Asp Val Asn Asp Trp Val Gly Pro Pro Asn Asn Asn Gly Val Ile Lys
        355                 360                 365
Glu Val Thr Ile Asn Pro Asp Thr Thr Cys Gly Asn Asp Trp Val Cys
    370                 375                 380
Glu His Arg Trp Arg Gln Ile Arg Asn Met Val Ile Phe Arg Asn Val
385                 390                 395                 400
Val Asp Gly Gln Pro Phe Thr Asn Trp Tyr Asp Asn Gly Ser Asn Gln
                405                 410                 415
Val Ala Phe Gly Arg Gly Asn Arg Gly Phe Ile Val Phe Asn Asn Asp
            420                 425                 430
Asp Trp Ser Phe Ser Leu Thr Leu Gln Thr Gly Leu Pro Ala Gly Thr
        435                 440                 445
Tyr Cys Asp Val Ile Ser Gly Asp Lys Ile Asn Gly Asn Cys Thr Gly
    450                 455                 460
Ile Lys Ile Tyr Val Ser Asp Asp Gly Lys Ala His Phe Ser Ile Ser
465                 470                 475                 480
Asn Ser Ala Glu Asp Pro Phe Ile Ala Ile His Ala Glu Ser Lys Leu
                485                 490                 495
```

<210> 33
<211> 370
<212> PRT
<213> Trichoderma reesei

<400> 33

```
Gln Pro Gly Thr Ser Thr Pro Glu Val His Pro Lys Leu Thr Thr Tyr
1               5                   10                  15
Lys Cys Thr Lys Ser Gly Gly Cys Val Ala Gln Asp Thr Ser Val Val
            20                  25                  30
Leu Asp Trp Asn Tyr Arg Trp Met His Asp Ala Asn Tyr Asn Ser Cys
        35                  40                  45
Thr Val Asn Gly Gly Val Asn Thr Thr Leu Cys Pro Asp Glu Ala Thr
```

```
    50                  55                  60
Cys Gly Lys Asn Cys Phe Ile Glu Gly Val Asp Tyr Ala Ala Ser Gly
65                  70                  75                  80
Val Thr Thr Ser Gly Ser Ser Leu Thr Met Asn Gln Tyr Met Pro Ser
                85                  90                  95
Ser Ser Gly Gly Tyr Ser Ser Val Ser Pro Arg Leu Tyr Leu Leu Asp
            100                 105                 110
Ser Asp Gly Glu Tyr Val Met Leu Lys Leu Asn Gly Gln Glu Leu Ser
        115                 120                 125
Phe Asp Val Asp Leu Ser Ala Leu Pro Cys Gly Glu Asn Gly Ser Leu
    130                 135                 140
Tyr Leu Ser Gln Met Asp Glu Asn Gly Gly Ala Asn Gln Tyr Asn Thr
145                 150                 155                 160
Ala Gly Ala Asn Tyr Gly Ser Gly Tyr Cys Asp Ala Gln Cys Pro Val
                165                 170                 175
Gln Thr Trp Arg Asn Gly Thr Leu Asn Thr Ser His Gln Gly Phe Cys
            180                 185                 190
Cys Asn Glu Met Asp Ile Leu Glu Gly Asn Ser Arg Ala Asn Ala Leu
        195                 200                 205
Thr Pro His Ser Cys Thr Ala Thr Ala Cys Asp Ser Ala Gly Cys Gly
    210                 215                 220
Phe Asn Pro Tyr Gly Ser Gly Tyr Lys Ser Tyr Tyr Gly Pro Gly Asp
225                 230                 235                 240
Thr Val Asp Thr Ser Lys Thr Phe Thr Ile Ile Thr Gln Phe Asn Thr
                245                 250                 255
Asp Asn Gly Ser Pro Ser Gly Asn Leu Val Ser Ile Thr Arg Lys Tyr
            260                 265                 270
Gln Gln Asn Gly Val Asp Ile Pro Ser Ala Gln Pro Gly Gly Asp Thr
        275                 280                 285
Ile Ser Ser Cys Pro Ser Ala Ser Ala Tyr Gly Gly Leu Ala Thr Met
    290                 295                 300
Gly Lys Ala Leu Ser Ser Gly Met Val Leu Val Phe Ser Ile Trp Asn
305                 310                 315                 320
Asp Asn Ser Gln Tyr Met Asn Trp Leu Asp Ser Gly Asn Ala Gly Pro
                325                 330                 335
Cys Ser Ser Thr Glu Gly Asn Pro Ser Asn Ile Leu Ala Asn Asn Pro
            340                 345                 350
Asn Thr His Val Val Phe Ser Asn Ile Arg Trp Gly Asp Ile Gly Ser
        355                 360                 365
Thr Thr
    370
```

<210> 34
<211> 223
<212> PRT
<213> Aspergillus niger

<400> 34

```
Gln Thr Met Cys Ser Gln Tyr Asp Ser Ala Ser Ser Pro Pro Tyr Ser
1               5                   10                  15
Val Asn Gln Asn Leu Trp Gly Glu Tyr Gln Gly Thr Gly Ser Gln Cys
            20                  25                  30
Val Tyr Val Asp Lys Leu Ser Ser Ser Gly Ala Ser Trp His Thr Glu
        35                  40                  45
Trp Thr Trp Ser Gly Gly Glu Gly Thr Val Lys Ser Tyr Ser Asn Ser
    50                  55                  60
Gly Val Thr Phe Asn Lys Lys Leu Val Ser Asp Val Ser Ser Ile Pro
65                  70                  75                  80
Thr Ser Val Glu Trp Lys Gln Asp Asn Thr Asn Val Asn Ala Asp Val
                85                  90                  95
Ala Tyr Asp Leu Phe Thr Ala Ala Asn Val Asp His Ala Thr Ser Ser
            100                 105                 110
```

```
Gly Asp Tyr Glu Leu Met Ile Trp Leu Ala Arg Tyr Gly Asn Ile Gln
        115                 120                 125
Pro Ile Gly Lys Gln Ile Ala Thr Ala Thr Val Gly Gly Lys Ser Trp
    130                 135                 140
Glu Val Trp Tyr Gly Ser Thr Thr Gln Ala Gly Ala Glu Gln Arg Thr
145                 150                 155                 160
Tyr Ser Phe Val Ser Glu Ser Pro Ile Asn Ser Tyr Ser Gly Asp Ile
                165                 170                 175
Asn Ala Phe Phe Ser Tyr Leu Thr Gln Asn Gln Gly Phe Pro Ala Ser
            180                 185                 190
Ser Gln Tyr Leu Ile Asn Leu Gln Phe Gly Thr Glu Ala Phe Thr Gly
        195                 200                 205
Gly Pro Ala Thr Phe Thr Val Asp Asn Trp Thr Ala Ser Val Asn
    210                 215                 220
```

<210> 35
<211> 184
<212> PRT
<213> Aspergillus niger

<400> 35

```
Ser Ala Gly Ile Asn Tyr Val Gln Asn Tyr Asn Gly Asn Leu Gly Asp
1               5                   10                  15
Phe Thr Tyr Asp Glu Ser Ala Gly Thr Phe Ser Met Tyr Trp Glu Asp
            20                  25                  30
Gly Val Ser Ser Asp Phe Val Val Gly Leu Gly Trp Thr Thr Gly Ser
        35                  40                  45
Ser Asn Ala Ile Thr Tyr Ser Ala Glu Tyr Ser Ala Ser Gly Ser Ala
    50                  55                  60
Ser Tyr Leu Ala Val Tyr Gly Trp Val Asn Tyr Pro Gln Ala Glu Tyr
65                  70                  75                  80
Tyr Ile Val Glu Asp Tyr Gly Asp Tyr Asn Pro Cys Ser Ser Ala Thr
                85                  90                  95
Ser Leu Gly Thr Val Tyr Ser Asp Gly Ser Thr Tyr Gln Val Cys Thr
            100                 105                 110
Asp Thr Arg Thr Asn Glu Pro Ser Ile Thr Gly Thr Ser Thr Phe Thr
        115                 120                 125
Gln Tyr Phe Ser Val Arg Glu Ser Thr Arg Thr Ser Gly Thr Val Thr
    130                 135                 140
Val Ala Asn His Phe Asn Phe Trp Ala His His Gly Phe Gly Asn Ser
145                 150                 155                 160
Asp Phe Asn Tyr Gln Val Val Ala Val Glu Ala Trp Ser Gly Ala Gly
                165                 170                 175
Ser Ala Ser Val Thr Ile Ser Ser
            180
```

<210> 36
<211> 313
<212> PRT
<213> Streptomyces lividans

<400> 36

```
Ala Glu Ser Thr Leu Gly Ala Ala Ala Ala Gln Ser Gly Arg Tyr Phe
1               5                   10                  15
Gly Thr Ala Ile Ala Ser Gly Arg Leu Ser Asp Ser Thr Tyr Thr Ser
            20                  25                  30
Ile Ala Gly Arg Glu Phe Asn Met Val Thr Ala Glu Asn Glu Met Lys
        35                  40                  45
Ile Asp Ala Thr Glu Pro Gln Arg Gly Gln Phe Asn Phe Ser Ser Ala
    50                  55                  60
Asp Arg Val Tyr Asn Trp Ala Val Gln Asn Gly Lys Gln Val Arg Gly
```

65 70 75 80
His Thr Leu Ala Trp His Ser Gln Gln Pro Gly Trp Met Gln Ser Leu
85 90 95
Ser Gly Ser Ala Leu Arg Gln Ala Met Ile Asp His Ile Asn Gly Val
100 105 110
Met Ala His Tyr Lys Gly Lys Ile Val Gln Trp Asp Val Val Asn Glu
115 120 125
Ala Phe Ala Asp Gly Ser Ser Gly Ala Arg Arg Asp Ser Asn Leu Gln
130 135 140
Arg Ser Gly Asn Asp Trp Ile Glu Val Ala Phe Arg Thr Ala Arg Ala
145 150 155 160
Ala Asp Pro Ser Ala Lys Leu Cys Tyr Asn Asp Tyr Asn Val Glu Asn
165 170 175
Trp Thr Trp Ala Lys Thr Gln Ala Met Tyr Asn Met Val Arg Asp Phe
180 185 190
Lys Gln Arg Gly Val Pro Ile Asp Cys Val Gly Phe Gln Ser His Phe
195 200 205
Asn Ser Gly Ser Pro Tyr Asn Ser Asn Phe Arg Thr Thr Leu Gln Asn
210 215 220
Phe Ala Ala Leu Gly Val Asp Val Ala Ile Thr Glu Leu Asp Ile Gln
225 230 235 240
Gly Ala Pro Ala Ser Thr Tyr Ala Asn Val Thr Asn Asp Cys Leu Ala
245 250 255
Val Ser Arg Cys Leu Gly Ile Thr Val Trp Gly Val Arg Asp Ser Asp
260 265 270
Ser Trp Arg Ser Glu Gln Thr Pro Leu Leu Phe Asn Asn Asp Gly Ser
275 280 285
Lys Lys Ala Ala Tyr Thr Ala Val Leu Asp Ala Leu Asn Gly Gly Ala
290 295 300
Ser Ser Glu Pro Pro Ala Asp Gly Gly
305 310

<210> 37
<211> 362
<212> PRT
<213> Aspergillus niger

<400> 37
Met His Ser Phe Ala Ser Leu Leu Ala Tyr Gly Leu Val Ala Gly Ala
1 5 10 15
Thr Phe Ala Ser Ala Ser Pro Ile Glu Ala Arg Asp Ser Cys Thr Phe
20 25 30
Thr Thr Ala Ala Ala Ala Lys Ala Gly Lys Ala Lys Cys Ser Thr Ile
35 40 45
Thr Leu Asn Asn Ile Glu Val Pro Ala Gly Thr Thr Leu Asp Leu Thr
50 55 60
Gly Leu Thr Ser Gly Thr Lys Val Ile Phe Glu Gly Thr Thr Thr Phe
65 70 75 80
Gln Tyr Glu Glu Trp Ala Gly Pro Leu Ile Ser Met Ser Gly Glu His
85 90 95
Ile Thr Val Thr Gly Ala Ser Gly His Leu Ile Asn Cys Asp Gly Ala
100 105 110
Arg Trp Trp Asp Gly Lys Gly Thr Ser Gly Lys Lys Lys Pro Lys Phe
115 120 125
Phe Tyr Ala His Gly Leu Asp Ser Ser Ser Ile Thr Gly Leu Asn Ile
130 135 140
Lys Asn Thr Pro Leu Met Ala Phe Ser Val Gln Ala Asn Asp Ile Thr
145 150 155 160
Phe Thr Asp Val Thr Ile Asn Asn Ala Asp Gly Asp Thr Gln Gly Gly
165 170 175
His Asn Thr Asp Ala Phe Asp Val Gly Asn Ser Val Gly Val Asn Ile
180 185 190

```
Ile Lys Pro Trp Val His Asn Gln Asp Asp Cys Leu Ala Val Asn Ser
        195                 200                 205
Gly Glu Asn Ile Trp Phe Thr Gly Gly Thr Cys Ile Gly Gly His Gly
    210                 215                 220
Leu Ser Ile Gly Ser Val Gly Asp Arg Ser Asn Asn Val Val Lys Asn
225                 230                 235                 240
Val Thr Ile Glu His Ser Thr Val Ser Asn Ser Glu Asn Ala Val Arg
                245                 250                 255
Ile Lys Thr Ile Ser Gly Ala Thr Gly Ser Val Ser Glu Ile Thr Tyr
            260                 265                 270
Ser Asn Ile Val Met Ser Gly Ile Ser Asp Tyr Gly Val Val Ile Gln
        275                 280                 285
Gln Asp Tyr Glu Asp Gly Lys Pro Thr Gly Lys Pro Thr Asn Gly Val
    290                 295                 300
Thr Ile Gln Asp Val Lys Leu Glu Ser Val Thr Gly Ser Val Asp Ser
305                 310                 315                 320
Gly Ala Thr Glu Ile Tyr Leu Leu Cys Gly Ser Gly Ser Cys Ser Asp
                325                 330                 335
Trp Thr Trp Asp Asp Val Lys Val Thr Gly Gly Lys Lys Ser Thr Ala
            340                 345                 350
Cys Lys Asn Phe Pro Ser Val Ala Ser Cys
        355                 360
```

<210> 38
<211> 383
<212> PRT
<213> Pseudomonas cellulosa

<400> 38

```
Arg Ala Asp Val Lys Pro Val Thr Val Lys Leu Val Asp Ser Gln Ala
1               5                   10                  15
Thr Met Glu Thr Arg Ser Leu Phe Ala Phe Met Gln Glu Gln Arg Arg
            20                  25                  30
His Ser Ile Met Phe Gly His Gln His Glu Thr Thr Gln Gly Leu Thr
        35                  40                  45
Ile Thr Arg Thr Asp Gly Thr Gln Ser Asp Thr Phe Asn Ala Val Gly
    50                  55                  60
Asp Phe Ala Ala Val Tyr Gly Trp Asp Thr Leu Ser Ile Val Ala Pro
65                  70                  75                  80
Lys Ala Glu Gly Asp Ile Val Ala Gln Val Lys Lys Ala Tyr Ala Arg
                85                  90                  95
Gly Gly Ile Ile Thr Val Ser Ser His Phe Asp Asn Pro Lys Thr Asp
            100                 105                 110
Thr Gln Lys Gly Val Trp Pro Val Gly Thr Ser Trp Asp Gln Thr Pro
        115                 120                 125
Ala Val Val Asp Ser Leu Pro Gly Gly Ala Tyr Asn Pro Val Leu Asn
    130                 135                 140
Gly Tyr Leu Asp Gln Val Ala Glu Trp Ala Asn Asn Leu Lys Asp Glu
145                 150                 155                 160
Gln Gly Arg Leu Ile Pro Val Ile Phe Arg Leu Tyr His Ala Asn Thr
                165                 170                 175
Gly Ser Trp Phe Trp Trp Gly Asp Lys Gln Ser Thr Pro Glu Gln Tyr
            180                 185                 190
Lys Gln Leu Phe Arg Tyr Ser Val Glu Tyr Leu Arg Asp Val Lys Gly
        195                 200                 205
Val Arg Asn Phe Leu Tyr Ala Tyr Ser Pro Asn Asn Phe Trp Asp Val
    210                 215                 220
Thr Glu Ala Asn Tyr Leu Glu Arg Tyr Pro Gly Asp Glu Trp Val Asp
225                 230                 235                 240
Val Leu Gly Phe Asp Thr Tyr Gly Pro Val Ala Asp Asn Ala Asp Trp
                245                 250                 255
Phe Arg Asn Val Val Ala Asn Ala Ala Leu Val Ala Arg Met Ala Glu
```

```
            260                 265                 270
Ala Arg Gly Lys Ile Pro Val Ile Ser Glu Ile Gly Ile Arg Ala Pro
        275                 280                 285
Asp Ile Glu Ala Gly Leu Tyr Asp Asn Gln Trp Tyr Arg Lys Leu Ile
    290                 295                 300
Ser Gly Leu Lys Ala Asp Pro Asp Ala Arg Glu Ile Ala Phe Leu Leu
305                 310                 315                 320
Val Trp Arg Asn Ala Pro Gln Gly Val Pro Gly Pro Asn Gly Thr Gln
                325                 330                 335
Val Pro His Tyr Trp Val Pro Ala Asn Arg Pro Glu Asn Ile Asn Asn
            340                 345                 350
Gly Thr Leu Glu Asp Phe Gln Ala Phe Tyr Ala Asp Glu Phe Thr Ala
        355                 360                 365
Phe Asn Arg Asp Ile Glu Gln Val Tyr Gln Arg Pro Thr Leu Ile
    370                 375                 380


<210>  39
<211>  419
<212>  PRT
<213>  Bacillus circulans

<400>  39
Leu Gln Pro Ala Thr Ala Glu Ala Ala Asp Ser Tyr Lys Ile Val Gly
1               5                   10                  15
Tyr Tyr Pro Ser Trp Ala Ala Tyr Gly Arg Asn Tyr Asn Val Ala Asp
            20                  25                  30
Ile Asp Pro Thr Lys Val Thr His Ile Asn Tyr Ala Phe Ala Asp Ile
        35                  40                  45
Cys Trp Asn Gly Ile His Gly Asn Pro Asp Pro Ser Gly Pro Asn Pro
    50                  55                  60
Val Thr Trp Thr Cys Gln Asn Glu Lys Ser Gln Thr Ile Asn Val Pro
65                  70                  75                  80
Asn Gly Thr Ile Val Leu Gly Asp Pro Trp Ile Asp Thr Gly Lys Thr
                85                  90                  95
Phe Ala Gly Asp Thr Trp Asp Gln Pro Ile Ala Gly Asn Ile Asn Gln
            100                 105                 110
Leu Asn Lys Leu Lys Gln Thr Asn Pro Asn Leu Lys Thr Ile Ile Ser
        115                 120                 125
Val Gly Gly Trp Thr Trp Ser Asn Arg Phe Ser Asp Val Ala Ala Thr
    130                 135                 140
Ala Ala Thr Arg Glu Val Phe Ala Asn Ser Ala Val Asp Phe Leu Arg
145                 150                 155                 160
Lys Tyr Asn Phe Asp Gly Val Asp Leu Asp Trp Glu Tyr Pro Val Ser
                165                 170                 175
Gly Gly Leu Asp Gly Asn Ser Lys Arg Pro Glu Asp Lys Gln Asn Tyr
            180                 185                 190
Thr Leu Leu Leu Ser Lys Ile Arg Glu Lys Leu Asp Ala Ala Gly Ala
        195                 200                 205
Val Asp Gly Lys Lys Tyr Leu Leu Thr Ile Ala Ser Gly Ala Ser Ala
    210                 215                 220
Thr Tyr Ala Ala Asn Thr Glu Leu Ala Lys Ile Ala Ala Ile Val Asp
225                 230                 235                 240
Trp Ile Asn Ile Met Thr Tyr Asp Phe Asn Gly Ala Trp Gln Lys Ile
                245                 250                 255
Ser Ala His Asn Ala Pro Leu Asn Tyr Asp Pro Ala Ala Ser Ala Ala
            260                 265                 270
Gly Val Pro Asp Ala Asn Thr Phe Asn Val Ala Ala Gly Ala Gln Gly
        275                 280                 285
His Leu Asp Ala Gly Val Pro Ala Ala Lys Leu Val Leu Gly Val Pro
    290                 295                 300
Phe Tyr Gly Arg Gly Trp Asp Gly Cys Ala Gln Ala Gly Asn Gly Gln
305                 310                 315                 320
```

```
Tyr Gln Thr Cys Thr Gly Gly Ser Ser Val Gly Thr Trp Glu Ala Gly
                325                 330                 335
Ser Phe Asp Phe Tyr Asp Leu Glu Ala Asn Tyr Ile Asn Lys Asn Gly
            340                 345                 350
Tyr Thr Arg Tyr Trp Asn Asp Thr Ala Lys Val Pro Tyr Leu Tyr Asn
        355                 360                 365
Ala Ser Asn Lys Arg Phe Ile Ser Tyr Asp Asp Ala Glu Ser Val Gly
    370                 375                 380
Tyr Lys Thr Ala Tyr Ile Lys Ser Lys Gly Leu Gly Gly Ala Met Phe
385                 390                 395                 400
Trp Glu Leu Ser Gly Asp Arg Asn Lys Thr Leu Gln Asn Lys Leu Lys
                405                 410                 415
Ala Asp Leu
```

<210> 40
<211> 317
<212> PRT
<213> Candida antarctica

<400> 40

```
Leu Pro Ser Gly Ser Asp Pro Ala Phe Ser Gln Pro Lys Ser Val Leu
1               5                   10                  15
Asp Ala Gly Leu Thr Cys Gln Gly Ala Ser Pro Ser Ser Val Ser Lys
            20                  25                  30
Pro Ile Leu Leu Val Pro Gly Thr Gly Thr Thr Gly Pro Gln Ser Phe
        35                  40                  45
Asp Ser Asn Trp Ile Pro Leu Ser Thr Gln Leu Gly Tyr Thr Pro Cys
    50                  55                  60
Trp Ile Ser Pro Pro Pro Phe Met Leu Asn Asp Thr Gln Val Asn Thr
65                  70                  75                  80
Glu Tyr Met Val Asn Ala Ile Thr Ala Leu Tyr Ala Gly Ser Gly Asn
                85                  90                  95
Asn Lys Leu Pro Val Leu Thr Trp Ser Gln Gly Gly Leu Val Ala Gln
            100                 105                 110
Trp Gly Leu Thr Phe Phe Pro Ser Ile Arg Ser Lys Val Asp Arg Leu
        115                 120                 125
Met Ala Phe Ala Pro Asp Tyr Lys Gly Thr Val Leu Ala Gly Pro Leu
    130                 135                 140
Asp Ala Leu Ala Val Ser Ala Pro Ser Val Trp Gln Gln Thr Thr Gly
145                 150                 155                 160
Ser Ala Leu Thr Thr Ala Leu Arg Asn Ala Gly Gly Leu Thr Gln Ile
                165                 170                 175
Val Pro Thr Thr Asn Leu Tyr Ser Ala Thr Asp Glu Ile Val Gln Pro
            180                 185                 190
Gln Val Ser Asn Ser Pro Leu Asp Ser Ser Tyr Leu Phe Asn Gly Lys
        195                 200                 205
Asn Val Gln Ala Gln Ala Val Cys Gly Pro Leu Phe Val Ile Asp His
    210                 215                 220
Ala Gly Ser Leu Thr Ser Gln Phe Ser Tyr Val Val Gly Arg Ser Ala
225                 230                 235                 240
Leu Arg Ser Thr Thr Gly Gln Ala Arg Ser Ala Asp Tyr Gly Ile Thr
                245                 250                 255
Asp Cys Asn Pro Leu Pro Ala Asn Asp Leu Thr Pro Glu Gln Lys Val
            260                 265                 270
Ala Ala Ala Ala Leu Leu Ala Pro Ala Ala Ala Ala Ile Val Ala Gly
        275                 280                 285
Pro Lys Gln Asn Cys Glu Pro Asp Leu Met Pro Tyr Ala Arg Pro Phe
    290                 295                 300
Ala Val Gly Lys Arg Thr Cys Ser Gly Ile Val Thr Pro
305                 310                 315
```

<210> 41
<211> 434
<212> PRT
<213> artificial sequence
<220>
<223> chimera of guinea pig and homo sapiens (human= approx. last 30 amino acids)

<400> 41

```
Ala Glu Val Cys Tyr Ser His Leu Gly Cys Phe Ser Asp Glu Lys Pro
1               5                  10                  15
Trp Ala Gly Thr Ser Gln Arg Pro Ile Lys Ser Leu Pro Ser Asp Pro
            20                  25                  30
Lys Lys Ile Asn Thr Arg Phe Leu Leu Tyr Thr Asn Glu Asn Gln Asn
        35                  40                  45
Ser Tyr Gln Leu Ile Thr Ala Thr Asp Ile Ala Thr Ile Lys Ala Ser
    50                  55                  60
Asn Phe Asn Leu Asn Arg Lys Thr Arg Phe Ile Ile His Gly Phe Thr
65                  70                  75                  80
Asp Ser Gly Glu Asn Ser Trp Leu Ser Asp Met Cys Lys Asn Met Phe
                85                  90                  95
Gln Val Glu Lys Val Asn Cys Ile Cys Val Asp Trp Lys Gly Gly Ser
            100                 105                 110
Lys Ala Gln Tyr Ser Gln Ala Ser Gln Asn Ile Arg Val Val Gly Ala
        115                 120                 125
Glu Val Ala Tyr Leu Val Gln Val Leu Ser Thr Ser Leu Asn Tyr Ala
    130                 135                 140
Pro Glu Asn Val His Ile Ile Gly His Ser Leu Gly Ala His Thr Ala
145                 150                 155                 160
Gly Glu Ala Gly Lys Arg Leu Asn Gly Leu Val Gly Arg Ile Thr Gly
                165                 170                 175
Leu Asp Pro Ala Glu Pro Tyr Phe Gln Asp Thr Pro Glu Glu Val Arg
            180                 185                 190
Leu Asp Pro Ser Asp Ala Lys Phe Val Asp Val Ile His Thr Asp Ile
        195                 200                 205
Ser Pro Ile Leu Pro Ser Leu Gly Phe Gly Met Ser Gln Lys Val Gly
    210                 215                 220
His Met Asp Phe Phe Pro Asn Gly Gly Lys Asp Met Pro Gly Cys Lys
225                 230                 235                 240
Thr Gly Ile Ser Cys Asn His His Arg Ser Ile Glu Tyr Tyr His Ser
                245                 250                 255
Ser Ile Leu Asn Pro Glu Gly Phe Leu Gly Tyr Pro Cys Ala Ser Tyr
            260                 265                 270
Asp Glu Phe Gln Glu Ser Gly Cys Phe Pro Cys Pro Ala Lys Gly Cys
        275                 280                 285
Pro Lys Met Gly His Phe Ala Asp Gln Tyr Pro Gly Lys Thr Asn Ala
    290                 295                 300
Val Glu Gln Thr Phe Phe Leu Asn Thr Gly Ala Ser Asp Asn Phe Thr
305                 310                 315                 320
Arg Trp Arg Tyr Lys Val Thr Val Thr Leu Ser Gly Glu Lys Asp Pro
                325                 330                 335
Ser Gly Asn Ile Asn Val Ala Leu Leu Gly Lys Asn Gly Asn Ser Ala
            340                 345                 350
Gln Tyr Gln Val Phe Lys Gly Thr Leu Lys Pro Asp Ala Ser Tyr Thr
        355                 360                 365
Asn Ser Ile Asp Val Glu Leu Asn Val Gly Thr Ile Gln Lys Val Thr
    370                 375                 380
Phe Leu Trp Lys Arg Ser Gly Ile Ser Val Ser Lys Pro Lys Met Gly
385                 390                 395                 400
Ala Ser Arg Ile Thr Val Gln Ser Gly Lys Asp Gly Thr Lys Tyr Asn
                405                 410                 415
Phe Cys Ser Ser Asp Ile Val Gln Glu Asn Val Glu Gln Thr Leu Ser
```

420 425 430
Pro Cys

<210> 42
<211> 471
<212> PRT
<213> Escherichia coli

<400> 42
Met Lys Gln Ser Thr Ile Ala Leu Ala Leu Leu Pro Leu Leu Phe Thr
1 5 10 15
Pro Val Thr Lys Ala Arg Thr Pro Glu Met Pro Val Leu Glu Asn Arg
20 25 30
Ala Ala Gln Gly Asp Ile Thr Ala Pro Gly Gly Ala Arg Arg Leu Thr
35 40 45
Gly Asp Gln Thr Ala Ala Leu Arg Asp Ser Leu Ser Asp Lys Pro Ala
50 55 60
Lys Asn Ile Ile Leu Leu Ile Gly Asp Gly Met Gly Asp Ser Glu Ile
65 70 75 80
Thr Ala Ala Arg Asn Tyr Ala Glu Gly Ala Gly Gly Phe Phe Lys Gly
85 90 95
Ile Asp Ala Leu Pro Leu Thr Gly Gln Tyr Thr His Tyr Ala Leu Asn
100 105 110
Lys Lys Thr Gly Lys Pro Asp Tyr Val Thr Asp Ser Ala Ala Ser Ala
115 120 125
Thr Ala Trp Ser Thr Gly Val Lys Thr Tyr Asn Gly Ala Leu Gly Val
130 135 140
Asp Ile His Glu Lys Asp His Pro Thr Ile Leu Glu Met Ala Lys Ala
145 150 155 160
Ala Gly Leu Ala Thr Gly Asn Val Ser Thr Ala Glu Leu Gln Asp Ala
165 170 175
Thr Pro Ala Ala Leu Val Ala His Val Thr Ser Arg Lys Cys Tyr Gly
180 185 190
Pro Ser Ala Thr Ser Glu Lys Cys Pro Gly Asn Ala Leu Glu Lys Gly
195 200 205
Gly Lys Gly Ser Ile Thr Glu Gln Leu Leu Asn Ala Arg Ala Asp Val
210 215 220
Thr Leu Gly Gly Gly Ala Lys Thr Phe Ala Glu Thr Ala Thr Ala Gly
225 230 235 240
Glu Trp Gln Gly Lys Thr Leu Arg Glu Gln Ala Gln Ala Arg Gly Tyr
245 250 255
Gln Leu Val Ser Asp Ala Ala Ser Leu Asn Ser Val Thr Glu Ala Asn
260 265 270
Gln Gln Lys Pro Leu Leu Gly Leu Phe Ala Asp Gly Asn Met Pro Val
275 280 285
Arg Trp Leu Gly Pro Lys Ala Thr Tyr His Gly Asn Ile Asp Lys Pro
290 295 300
Ala Val Thr Cys Thr Pro Asn Pro Gln Arg Asn Asp Ser Val Pro Thr
305 310 315 320
Leu Ala Gln Met Thr Asp Lys Ala Ile Glu Leu Leu Ser Lys Asn Glu
325 330 335
Lys Gly Phe Phe Leu Gln Val Glu Gly Ala Ser Ile Asp Lys Gln Asp
340 345 350
His Ala Ala Asn Pro Cys Gly Gln Ile Gly Glu Thr Val Asp Leu Asp
355 360 365
Glu Ala Val Gln Arg Ala Leu Glu Phe Ala Lys Lys Glu Gly Asn Thr
370 375 380
Leu Val Ile Val Thr Ala Asp His Ala His Ala Ser Gln Ile Val Ala
385 390 395 400
Pro Asp Thr Lys Ala Pro Gly Leu Thr Gln Ala Leu Asn Thr Lys Asp
405 410 415

```
Gly Ala Val Met Val Met Ser Tyr Gly Asn Ser Glu Glu Asp Ser Gln
            420                 425                 430
Glu His Thr Gly Ser Gln Leu Arg Ile Ala Ala Tyr Gly Pro His Ala
        435                 440                 445
Ala Asn Val Val Gly Leu Thr Asp Gln Thr Asp Leu Phe Tyr Thr Met
    450                 455                 460
Lys Ala Ala Leu Gly Leu Lys
465                 470
```

<210> 43
<211> 260
<212> PRT
<213> Bovine

<400> 43

```
Leu Lys Ile Ala Ala Phe Asn Ile Arg Thr Phe Gly Glu Thr Lys Met
1               5                   10                  15
Ser Asn Ala Thr Leu Ala Ser Tyr Ile Val Arg Ile Val Arg Arg Tyr
            20                  25                  30
Asp Ile Val Leu Ile Gln Glu Val Arg Asp Ser His Leu Val Ala Val
        35                  40                  45
Gly Lys Leu Leu Asp Tyr Leu Asn Gln Asp Asp Pro Asn Thr Tyr His
    50                  55                  60
Tyr Val Val Ser Glu Pro Leu Gly Arg Asn Ser Tyr Lys Glu Arg Tyr
65                  70                  75                  80
Leu Phe Leu Phe Arg Pro Asn Lys Val Ser Val Leu Asp Thr Tyr Gln
                85                  90                  95
Tyr Asp Asp Gly Cys Glu Ser Cys Gly Asn Asp Ser Phe Ser Arg Glu
            100                 105                 110
Pro Ala Val Val Lys Phe Ser Ser His Ser Thr Lys Val Lys Glu Phe
        115                 120                 125
Ala Ile Val Ala Leu His Ser Ala Pro Ser Asp Ala Val Ala Glu Ile
    130                 135                 140
Asn Ser Leu Tyr Asp Val Tyr Leu Asp Val Gln Gln Lys Trp His Leu
145                 150                 155                 160
Asn Asp Val Met Leu Met Gly Asp Phe Asn Ala Asp Cys Ser Tyr Val
                165                 170                 175
Thr Ser Ser Gln Trp Ser Ser Ile Arg Leu Arg Thr Ser Ser Thr Phe
            180                 185                 190
Gln Trp Leu Ile Pro Asp Ser Ala Asp Thr Thr Ala Thr Ser Thr Asn
        195                 200                 205
Cys Ala Tyr Asp Arg Ile Val Val Ala Gly Ser Leu Leu Gln Ser Ser
    210                 215                 220
Val Val Pro Gly Ser Ala Ala Pro Phe Asp Phe Gln Ala Ala Tyr Gly
225                 230                 235                 240
Leu Ser Asn Glu Met Ala Leu Ala Ile Ser Asp His Tyr Pro Val Glu
                245                 250                 255
Val Thr Leu Thr
            260
```

<210> 44
<211> 686
<212> PRT
<213> Bacillus circulans

<400> 44

```
Ala Pro Asp Thr Ser Val Ser Asn Lys Gln Asn Phe Ser Thr Asp Val
1               5                   10                  15
Ile Tyr Gln Ile Phe Thr Asp Arg Phe Ser Asp Gly Asn Pro Ala Asn
            20                  25                  30
Asn Pro Thr Gly Ala Ala Phe Asp Gly Thr Cys Thr Asn Leu Arg Leu
```

```
        35                  40                  45
Tyr Cys Gly Gly Asp Trp Gln Gly Ile Ile Asn Lys Ile Asn Asp Gly
    50                  55                  60
Tyr Leu Thr Gly Met Gly Val Thr Ala Ile Trp Ile Ser Gln Pro Val
65                  70                  75                  80
Glu Asn Ile Tyr Ser Ile Ile Asn Tyr Ser Gly Val Asn Asn Thr Ala
                85                  90                  95
Tyr His Gly Tyr Trp Ala Arg Asp Phe Lys Lys Thr Asn Pro Ala Tyr
            100                 105                 110
Gly Thr Ile Ala Asp Phe Gln Asn Leu Ile Ala Ala Ala His Ala Lys
        115                 120                 125
Asn Ile Lys Val Ile Ile Asp Phe Ala Pro Asn His Thr Ser Pro Ala
    130                 135                 140
Ser Ser Asp Gln Pro Ser Phe Ala Glu Asn Gly Arg Leu Tyr Asp Asn
145                 150                 155                 160
Gly Thr Leu Leu Gly Gly Tyr Thr Asn Asp Thr Gln Asn Leu Phe His
                165                 170                 175
His Asn Gly Gly Thr Asp Phe Ser Thr Thr Glu Asn Gly Ile Tyr Lys
            180                 185                 190
Asn Leu Tyr Asp Leu Ala Asp Leu Asn His Asn Asn Ser Thr Val Asp
        195                 200                 205
Val Tyr Leu Lys Asp Ala Ile Lys Met Trp Leu Asp Leu Gly Ile Asp
    210                 215                 220
Gly Ile Arg Met Asp Ala Val Lys His Met Pro Phe Gly Trp Gln Lys
225                 230                 235                 240
Ser Phe Met Ala Ala Val Asn Asn Tyr Lys Pro Val Phe Thr Phe Gly
                245                 250                 255
Glu Trp Phe Leu Gly Val Asn Glu Val Ser Pro Glu Asn His Lys Phe
            260                 265                 270
Ala Asn Glu Ser Gly Met Ser Leu Leu Asp Phe Arg Phe Ala Gln Lys
        275                 280                 285
Val Arg Gln Val Phe Arg Asp Asn Thr Asp Asn Met Tyr Gly Leu Lys
    290                 295                 300
Ala Met Leu Glu Gly Ser Ala Ala Asp Tyr Ala Gln Val Asp Asp Gln
305                 310                 315                 320
Val Thr Phe Ile Asp Asn His Asp Met Glu Arg Phe His Ala Ser Asn
                325                 330                 335
Ala Asn Arg Arg Lys Leu Glu Gln Ala Leu Ala Phe Thr Leu Thr Ser
            340                 345                 350
Arg Gly Val Pro Ala Ile Tyr Tyr Gly Thr Glu Gln Tyr Met Ser Gly
        355                 360                 365
Gly Thr Asp Pro Asp Asn Arg Ala Arg Ile Pro Ser Phe Ser Thr Ser
    370                 375                 380
Thr Thr Ala Tyr Gln Val Ile Gln Lys Leu Ala Pro Leu Arg Lys Cys
385                 390                 395                 400
Asn Pro Ala Ile Ala Tyr Gly Ser Thr Gln Glu Arg Trp Ile Asn Asn
                405                 410                 415
Asp Val Leu Ile Tyr Glu Arg Lys Phe Gly Ser Asn Val Ala Val Val
            420                 425                 430
Ala Val Asn Arg Asn Leu Asn Ala Pro Ala Ser Ile Ser Gly Leu Val
        435                 440                 445
Thr Ser Leu Pro Gln Gly Ser Tyr Asn Asp Val Leu Gly Gly Leu Leu
    450                 455                 460
Asn Gly Asn Thr Leu Ser Val Gly Ser Gly Gly Ala Ala Ser Asn Phe
465                 470                 475                 480
Thr Leu Ala Ala Gly Gly Thr Ala Val Trp Gln Tyr Thr Ala Ala Thr
                485                 490                 495
Ala Thr Pro Thr Ile Gly His Val Gly Pro Met Met Ala Lys Pro Gly
            500                 505                 510
Val Thr Ile Thr Ile Asp Gly Arg Gly Phe Gly Ser Ser Lys Gly Thr
        515                 520                 525
Val Tyr Phe Gly Thr Thr Ala Val Ser Gly Ala Asp Ile Thr Ser Trp
    530                 535                 540
```

```
Glu Asp Thr Gln Ile Lys Val Lys Ile Pro Ala Val Ala Gly Gly Asn
545                 550                 555                 560
Tyr Asn Ile Lys Val Ala Asn Ala Ala Gly Thr Ala Ser Asn Val Tyr
                565                 570                 575
Asp Asn Phe Glu Val Leu Ser Gly Asp Gln Val Ser Val Arg Phe Val
            580                 585                 590
Val Asn Asn Ala Thr Thr Ala Leu Gly Gln Asn Val Tyr Leu Thr Gly
        595                 600                 605
Ser Val Ser Glu Leu Gly Asn Trp Asp Pro Ala Lys Ala Ile Gly Pro
    610                 615                 620
Met Tyr Asn Gln Val Val Tyr Gln Tyr Pro Asn Trp Tyr Tyr Asp Val
625                 630                 635                 640
Ser Val Pro Ala Gly Lys Thr Ile Glu Phe Lys Phe Leu Lys Lys Gln
                645                 650                 655
Gly Ser Thr Val Thr Trp Glu Gly Gly Ser Asn His Thr Phe Thr Ala
            660                 665                 670
Pro Ser Ser Gly Thr Ala Thr Ile Asn Val Asn Trp Gln Pro
        675                 680                 685
```

<210> 45
<211> 404
<212> PRT
<213> Amycolatopsis orientalis

<400> 45

```
Met Arg Val Leu Ile Thr Gly Cys Gly Ser Arg Gly Asp Thr Glu Pro
1               5                   10                  15
Leu Val Ala Leu Ala Ala Arg Leu Arg Glu Leu Gly Ala Asp Ala Arg
            20                  25                  30
Met Cys Leu Pro Pro Asp Tyr Val Glu Arg Cys Ala Glu Val Gly Val
        35                  40                  45
Pro Met Val Pro Val Gly Arg Ala Val Arg Ala Gly Ala Arg Glu Pro
    50                  55                  60
Gly Glu Leu Pro Pro Gly Ala Ala Glu Val Val Thr Glu Val Val Ala
65                  70                  75                  80
Glu Trp Phe Asp Lys Val Pro Ala Ala Ile Glu Gly Cys Asp Ala Val
                85                  90                  95
Val Thr Thr Gly Leu Leu Pro Ala Ala Val Ala Val Arg Ser Met Ala
            100                 105                 110
Glu Lys Leu Gly Ile Pro Tyr Arg Tyr Thr Val Leu Ser Pro Asp His
        115                 120                 125
Leu Pro Ser Glu Gln Ser Gln Ala Glu Arg Asp Met Tyr Asn Gln Gly
    130                 135                 140
Ala Asp Arg Leu Phe Gly Asp Ala Val Asn Ser His Arg Ala Ser Ile
145                 150                 155                 160
Gly Leu Pro Pro Val Glu His Leu Tyr Asp Tyr Gly Tyr Thr Asp Gln
                165                 170                 175
Pro Trp Leu Ala Ala Asp Pro Val Leu Ser Pro Leu Arg Pro Thr Asp
            180                 185                 190
Leu Gly Thr Val Gln Thr Gly Ala Trp Ile Leu Pro Asp Glu Arg Pro
        195                 200                 205
Leu Ser Ala Glu Leu Glu Ala Phe Leu Ala Ala Gly Ser Thr Pro Val
    210                 215                 220
Tyr Val Gly Phe Gly Ser Ser Ser Arg Pro Ala Thr Ala Asp Ala Ala
225                 230                 235                 240
Lys Met Ala Ile Lys Ala Val Arg Ala Ser Gly Arg Arg Ile Val Leu
                245                 250                 255
Ser Arg Gly Trp Ala Asp Leu Val Leu Pro Asp Asp Gly Ala Asp Cys
            260                 265                 270
Phe Val Val Gly Glu Val Asn Leu Gln Glu Leu Phe Gly Arg Val Ala
        275                 280                 285
Ala Ala Ile His His Asp Ser Ala Gly Thr Thr Leu Leu Ala Met Arg
```

```
    290                 295                 300
Ala Gly Ile Pro Gln Ile Val Val Arg Arg Val Val Asp Asn Val Val
305                 310                 315                 320
Glu Gln Ala Tyr His Ala Asp Arg Val Ala Glu Leu Gly Val Gly Val
                325                 330                 335
Ala Val Asp Gly Pro Val Pro Thr Ile Asp Ser Leu Ser Ala Ala Leu
            340                 345                 350
Asp Thr Ala Leu Ala Pro Glu Ile Arg Ala Arg Ala Thr Thr Val Ala
        355                 360                 365
Asp Thr Ile Arg Ala Asp Gly Thr Thr Val Ala Ala Gln Leu Leu Phe
    370                 375                 380
Asp Ala Val Ser Leu Glu Lys Pro Thr Val Pro Ala Leu Glu His His
385                 390                 395                 400
His His His His
```

<210> 46
<211> 292
<212> PRT
<213> Pseudomonas sp.

<400> 46

```
Ser Ile Glu Arg Leu Gly Tyr Leu Gly Phe Ala Val Lys Asp Val Pro
1               5                   10                  15
Ala Trp Asp His Phe Leu Thr Lys Ser Val Gly Leu Met Ala Ala Gly
            20                  25                  30
Ser Ala Gly Asp Ala Ala Leu Tyr Arg Ala Asp Gln Arg Ala Trp Arg
        35                  40                  45
Ile Ala Val Gln Pro Gly Glu Leu Asp Asp Leu Ala Tyr Ala Gly Leu
    50                  55                  60
Glu Val Asp Asp Ala Ala Ala Leu Glu Arg Met Ala Asp Lys Leu Arg
65                  70                  75                  80
Gln Ala Gly Val Ala Phe Thr Arg Gly Asp Glu Ala Leu Met Gln Gln
                85                  90                  95
Arg Lys Val Met Gly Leu Leu Cys Leu Gln Asp Pro Phe Gly Leu Pro
            100                 105                 110
Leu Glu Ile Tyr Tyr Gly Pro Ala Glu Ile Phe His Glu Pro Phe Leu
        115                 120                 125
Pro Ser Ala Pro Val Ser Gly Phe Val Thr Gly Asp Gln Gly Ile Gly
    130                 135                 140
His Phe Val Arg Cys Val Pro Asp Thr Ala Lys Ala Met Ala Phe Tyr
145                 150                 155                 160
Thr Glu Val Leu Gly Phe Val Leu Ser Asp Ile Ile Asp Ile Gln Met
                165                 170                 175
Gly Pro Glu Thr Ser Val Pro Ala His Phe Leu His Cys Asn Gly Arg
            180                 185                 190
His His Thr Ile Ala Leu Ala Ala Phe Pro Ile Pro Lys Arg Ile His
        195                 200                 205
His Phe Met Leu Gln Ala Asn Thr Ile Asp Asp Val Gly Tyr Ala Phe
    210                 215                 220
Asp Arg Leu Asp Ala Ala Gly Arg Ile Thr Ser Leu Leu Gly Arg His
225                 230                 235                 240
Thr Asn Asp Gln Thr Leu Ser Phe Tyr Ala Asp Thr Pro Ser Pro Met
                245                 250                 255
Ile Glu Val Glu Phe Gly Trp Gly Pro Arg Thr Val Asp Ser Ser Trp
            260                 265                 270
Thr Val Ala Arg His Ser Arg Thr Ala Met Trp Gly His Lys Ser Val
        275                 280                 285
Arg Gly Gln Arg
    290
```

<210> 47
<211> 311
<212> PRT
<213> Acitenobacter sp.

<400> 47
Met Glu Val Lys Ile Phe Asn Thr Gln Asp Val Gln Asp Phe Leu Arg
1 5 10 15
Val Ala Ser Gly Leu Glu Gln Glu Gly Gly Asn Pro Arg Val Lys Gln
20 25 30
Ile Ile His Arg Val Leu Ser Asp Leu Tyr Lys Ala Ile Glu Asp Leu
35 40 45
Asn Ile Thr Ser Asp Glu Tyr Trp Ala Gly Val Ala Tyr Leu Asn Gln
50 55 60
Leu Gly Ala Asn Gln Glu Ala Gly Leu Leu Ser Pro Gly Leu Gly Phe
65 70 75 80
Asp His Tyr Leu Asp Met Arg Met Asp Ala Glu Asp Ala Ala Leu Gly
85 90 95
Ile Glu Asn Ala Thr Pro Arg Thr Ile Glu Gly Pro Leu Tyr Val Ala
100 105 110
Gly Ala Pro Glu Ser Val Gly Tyr Ala Arg Met Asp Asp Gly Ser Asp
115 120 125
Pro Asn Gly His Thr Leu Ile Leu His Gly Thr Ile Phe Asp Ala Asp
130 135 140
Gly Lys Pro Leu Pro Asn Ala Lys Val Glu Ile Trp His Ala Asn Thr
145 150 155 160
Lys Gly Phe Tyr Ser His Phe Asp Pro Thr Gly Glu Gln Gln Ala Phe
165 170 175
Asn Met Arg Arg Ser Ile Ile Thr Asp Glu Asn Gly Gln Tyr Arg Val
180 185 190
Arg Thr Ile Leu Pro Ala Gly Tyr Gly Cys Pro Pro Glu Gly Pro Thr
195 200 205
Gln Gln Leu Leu Asn Gln Leu Gly Arg His Gly Asn Arg Pro Ala His
210 215 220
Ile His Tyr Phe Val Ser Ala Asp Gly His Arg Lys Leu Thr Thr Gln
225 230 235 240
Ile Asn Val Ala Gly Asp Pro Tyr Thr Tyr Asp Asp Phe Ala Tyr Ala
245 250 255
Thr Arg Glu Gly Leu Val Val Asp Ala Val Glu His Thr Asp Pro Glu
260 265 270
Ala Ile Lys Ala Asn Asp Val Glu Gly Pro Phe Ala Glu Met Val Phe
275 280 285
Asp Leu Lys Leu Thr Arg Leu Val Asp Gly Val Asp Asn Gln Val Val
290 295 300
Asp Arg Pro Arg Leu Ala Val
305 310

<210> 48
<211> 414
<212> PRT
<213> Pseudomonas putida

<400> 48
Thr Thr Glu Thr Ile Gln Ser Asn Ala Asn Leu Ala Pro Leu Pro Pro
1 5 10 15
His Val Pro Glu His Leu Val Phe Asp Phe Asp Met Tyr Asn Pro Ser
20 25 30
Asn Leu Ser Ala Gly Val Gln Glu Ala Trp Ala Val Leu Gln Glu Ser
35 40 45
Asn Val Pro Asp Leu Val Trp Thr Arg Cys Asn Gly Gly His Trp Ile
50 55 60
Ala Thr Arg Gly Gln Leu Ile Arg Glu Ala Tyr Glu Asp Tyr Arg His

65 70 75 80
Phe Ser Ser Glu Cys Pro Phe Ile Pro Arg Glu Ala Gly Glu Ala Tyr
85 90 95
Asp Phe Ile Pro Thr Ser Met Asp Pro Pro Glu Gln Arg Gln Phe Arg
100 105 110
Ala Leu Ala Asn Gln Val Val Gly Met Pro Val Val Asp Lys Leu Glu
115 120 125
Asn Arg Ile Gln Glu Leu Ala Cys Ser Leu Ile Glu Ser Leu Arg Pro
130 135 140
Gln Gly Gln Cys Asn Phe Thr Glu Asp Tyr Ala Glu Pro Phe Pro Ile
145 150 155 160
Arg Ile Phe Met Leu Leu Ala Gly Leu Pro Glu Glu Asp Ile Pro His
165 170 175
Leu Lys Tyr Leu Thr Asp Gln Met Thr Arg Pro Asp Gly Ser Met Thr
180 185 190
Phe Ala Glu Ala Lys Glu Ala Leu Tyr Asp Tyr Leu Ile Pro Ile Ile
195 200 205
Glu Gln Arg Arg Gln Lys Pro Gly Thr Asp Ala Ile Ser Ile Val Ala
210 215 220
Asn Gly Gln Val Asn Gly Arg Pro Ile Thr Ser Asp Glu Ala Lys Arg
225 230 235 240
Met Cys Gly Leu Leu Leu Val Gly Gly Leu Asp Thr Val Val Asn Phe
245 250 255
Leu Ser Phe Ser Met Glu Phe Leu Ala Lys Ser Pro Glu His Arg Gln
260 265 270
Glu Leu Ile Gln Arg Pro Glu Arg Ile Pro Ala Ala Cys Glu Glu Leu
275 280 285
Leu Arg Arg Phe Ser Leu Val Ala Asp Gly Arg Ile Leu Thr Ser Asp
290 295 300
Tyr Glu Phe His Gly Val Gln Leu Lys Lys Gly Asp Gln Ile Leu Leu
305 310 315 320
Pro Gln Met Leu Ser Gly Leu Asp Glu Arg Glu Asn Ala Cys Pro Met
325 330 335
His Val Asp Phe Ser Arg Gln Lys Val Ser His Thr Thr Phe Gly His
340 345 350
Gly Ser His Leu Cys Leu Gly Gln His Leu Ala Arg Arg Glu Ile Ile
355 360 365
Val Thr Leu Lys Glu Trp Leu Thr Arg Ile Pro Asp Phe Ser Ile Ala
370 375 380
Pro Gly Ala Gln Ile Gln His Lys Ser Gly Ile Val Ser Gly Val Gln
385 390 395 400
Ala Leu Pro Leu Val Trp Asp Pro Ala Thr Thr Lys Ala Val
405 410

<210> 49
<211> 374
<212> PRT
<213> Equus caballus

<400> 49
Ser Thr Ala Gly Lys Val Ile Lys Cys Lys Ala Ala Val Leu Trp Glu
1 5 10 15
Glu Lys Lys Pro Phe Ser Ile Glu Glu Val Glu Val Ala Pro Pro Lys
20 25 30
Ala His Glu Val Arg Ile Lys Met Val Ala Thr Gly Ile Cys Arg Ser
35 40 45
Asp Asp His Val Val Ser Gly Thr Leu Val Thr Pro Leu Pro Val Ile
50 55 60
Ala Gly His Glu Ala Ala Gly Ile Val Glu Ser Ile Gly Glu Gly Val
65 70 75 80
Thr Thr Val Arg Pro Gly Asp Lys Val Ile Pro Leu Phe Thr Pro Gln
85 90 95

```
Cys Gly Lys Cys Arg Val Cys Lys His Pro Glu Gly Asn Phe Cys Leu
            100                 105                 110
Lys Asn Asp Leu Ser Met Pro Arg Gly Thr Met Gln Asp Gly Thr Ser
        115                 120                 125
Arg Phe Thr Cys Arg Gly Lys Pro Ile His His Phe Leu Gly Thr Ser
    130                 135                 140
Thr Phe Ser Gln Tyr Thr Val Val Asp Glu Ile Ser Val Ala Lys Ile
145                 150                 155                 160
Asp Ala Ala Ser Pro Leu Glu Lys Val Cys Leu Ile Gly Cys Gly Phe
                165                 170                 175
Ser Thr Gly Tyr Gly Ser Ala Val Lys Val Ala Lys Val Thr Gln Gly
            180                 185                 190
Ser Thr Cys Ala Val Phe Gly Leu Gly Gly Val Gly Leu Ser Val Ile
        195                 200                 205
Met Gly Cys Lys Ala Ala Gly Ala Ala Arg Ile Ile Gly Val Asp Ile
    210                 215                 220
Asn Lys Asp Lys Phe Ala Lys Ala Lys Glu Val Gly Ala Thr Glu Cys
225                 230                 235                 240
Val Asn Pro Gln Asp Tyr Lys Lys Pro Ile Gln Glu Val Leu Thr Glu
                245                 250                 255
Met Ser Asn Gly Gly Val Asp Phe Ser Phe Glu Val Ile Gly Arg Leu
            260                 265                 270
Asp Thr Met Val Thr Ala Leu Ser Cys Cys Gln Glu Ala Tyr Gly Val
        275                 280                 285
Ser Val Ile Val Gly Val Pro Pro Asp Ser Gln Asn Leu Ser Met Asn
    290                 295                 300
Pro Met Leu Leu Leu Ser Gly Arg Thr Trp Lys Gly Ala Ile Phe Gly
305                 310                 315                 320
Gly Phe Lys Ser Lys Asp Ser Val Pro Lys Leu Val Ala Asp Phe Met
                325                 330                 335
Ala Lys Lys Phe Ala Leu Asp Pro Leu Ile Thr His Val Leu Pro Phe
            340                 345                 350
Glu Lys Ile Asn Glu Gly Phe Asp Leu Leu Arg Ser Gly Glu Ser Ile
        355                 360                 365
Arg Thr Ile Leu Thr Phe
    370
```

<210> 50
<211> 297
<212> PRT
<213> Escherichia coli

<400> 50

```
Met Ala Thr Asn Leu Arg Gly Val Met Ala Ala Leu Leu Thr Pro Phe
1               5                   10                  15
Asp Gln Gln Gln Ala Leu Asp Lys Ala Ser Leu Arg Arg Leu Val Gln
            20                  25                  30
Phe Asn Ile Gln Gln Gly Ile Asp Gly Leu Tyr Val Gly Gly Ser Thr
        35                  40                  45
Gly Glu Ala Phe Val Gln Ser Leu Ser Glu Arg Glu Gln Val Leu Glu
    50                  55                  60
Ile Val Ala Glu Glu Gly Lys Gly Lys Ile Lys Leu Ile Ala His Val
65                  70                  75                  80
Gly Cys Val Thr Thr Ala Glu Ser Gln Gln Leu Ala Ala Ser Ala Lys
                85                  90                  95
Arg Tyr Gly Phe Asp Ala Val Ser Ala Val Thr Pro Phe Tyr Tyr Pro
            100                 105                 110
Phe Ser Phe Glu Glu His Cys Asp His Tyr Arg Ala Ile Ile Asp Ser
        115                 120                 125
Ala Asp Gly Leu Pro Met Val Val Tyr Asn Ile Pro Ala Leu Ser Gly
    130                 135                 140
Val Lys Leu Thr Leu Asp Gln Ile Asn Thr Leu Val Thr Leu Pro Gly
```

```
145                 150                 155                 160
Val Gly Ala Leu Lys Gln Thr Ser Gly Asp Leu Tyr Gln Met Glu Gln
                165                 170                 175
Ile Arg Arg Glu His Pro Asp Leu Val Leu Tyr Asn Gly Tyr Asp Glu
            180                 185                 190
Ile Phe Ala Ser Gly Leu Leu Ala Gly Ala Asp Gly Gly Ile Gly Ser
        195                 200                 205
Thr Tyr Asn Ile Met Gly Trp Arg Tyr Gln Gly Ile Val Lys Ala Leu
    210                 215                 220
Lys Glu Gly Asp Ile Gln Thr Ala Gln Lys Leu Gln Thr Glu Cys Asn
225                 230                 235                 240
Lys Val Ile Asp Leu Leu Ile Lys Thr Gly Val Phe Arg Gly Leu Lys
                245                 250                 255
Thr Val Leu His Tyr Met Asp Val Val Ser Val Pro Leu Cys Arg Lys
            260                 265                 270
Pro Phe Gly Pro Val Asp Glu Lys Tyr Leu Pro Glu Leu Lys Ala Leu
        275                 280                 285
Ala Gln Gln Leu Met Gln Glu Arg Gly
    290                 295
```

<210> 51
<211> 268
<212> PRT
<213> Salmonella typhimurium

<400> 51

```
Met Glu Arg Tyr Glu Asn Leu Phe Ala Gln Leu Asn Asp Arg Arg Glu
1               5                  10                  15
Gly Ala Phe Val Pro Phe Val Thr Leu Gly Asp Pro Gly Ile Glu Gln
            20                  25                  30
Ser Leu Lys Ile Ile Asp Thr Leu Ile Asp Ala Gly Ala Asp Ala Leu
        35                  40                  45
Glu Leu Gly Val Pro Phe Ser Asp Pro Leu Ala Asp Gly Pro Thr Ile
    50                  55                  60
Gln Asn Ala Asn Leu Arg Ala Phe Ala Ala Gly Val Thr Pro Ala Gln
65                  70                  75                  80
Cys Phe Glu Met Leu Ala Leu Ile Arg Glu Lys His Pro Thr Ile Pro
                85                  90                  95
Ile Gly Leu Leu Met Tyr Ala Asn Leu Val Phe Asn Asn Gly Ile Asp
            100                 105                 110
Ala Phe Tyr Ala Arg Cys Glu Gln Val Gly Val Asp Ser Val Leu Val
        115                 120                 125
Ala Asp Val Pro Val Glu Glu Ser Ala Pro Phe Arg Gln Ala Ala Leu
    130                 135                 140
Arg His Asn Ile Ala Pro Ile Phe Ile Cys Pro Pro Asn Ala Asp Asp
145                 150                 155                 160
Asp Leu Leu Arg Gln Val Ala Ser Tyr Gly Arg Gly Tyr Thr Tyr Leu
                165                 170                 175
Leu Ser Arg Ser Gly Val Thr Gly Ala Glu Asn Arg Gly Ala Leu Pro
            180                 185                 190
Leu His His Leu Ile Glu Lys Leu Lys Glu Tyr His Ala Ala Pro Ala
        195                 200                 205
Leu Gln Gly Phe Gly Ile Ser Ser Pro Glu Gln Val Ser Ala Ala Val
    210                 215                 220
Arg Ala Gly Ala Ala Gly Ala Ile Ser Gly Ser Ala Ile Val Lys Ile
225                 230                 235                 240
Ile Glu Lys Asn Leu Ala Ser Pro Lys Gln Met Leu Ala Glu Leu Arg
                245                 250                 255
Ser Phe Val Ser Ala Met Lys Ala Ala Ser Arg Ala
            260                 265
```

<210> 52
<211> 393
<212> PRT
<213> Actinoplanes missouriensis

<400> 52
Ser Val Gln Ala Thr Arg Glu Asp Lys Phe Ser Phe Gly Leu Trp Thr
1 5 10 15
Val Gly Trp Gln Ala Arg Asp Ala Phe Gly Asp Ala Thr Arg Thr Ala
20 25 30
Leu Asp Pro Val Glu Ala Val His Lys Leu Ala Glu Ile Gly Ala Tyr
35 40 45
Gly Ile Thr Phe His Asp Asp Asp Leu Val Pro Phe Gly Ser Asp Ala
50 55 60
Gln Thr Arg Asp Gly Ile Ile Ala Gly Phe Lys Lys Ala Leu Asp Glu
65 70 75 80
Thr Gly Leu Ile Val Pro Met Val Thr Thr Asn Leu Phe Thr His Pro
85 90 95
Val Phe Lys Asp Gly Gly Phe Thr Ser Asn Asp Arg Ser Val Arg Arg
100 105 110
Tyr Ala Ile Arg Lys Val Leu Arg Gln Met Asp Leu Gly Ala Glu Leu
115 120 125
Gly Ala Lys Thr Leu Val Leu Trp Gly Gly Arg Glu Gly Ala Glu Tyr
130 135 140
Asp Ser Ala Lys Asp Val Ser Ala Ala Leu Asp Arg Tyr Arg Glu Ala
145 150 155 160
Leu Asn Leu Leu Ala Gln Tyr Ser Glu Asp Arg Gly Tyr Gly Leu Arg
165 170 175
Phe Ala Ile Glu Pro Lys Pro Asn Glu Pro Arg Gly Asp Ile Leu Leu
180 185 190
Pro Thr Ala Gly His Ala Ile Ala Phe Val Gln Glu Leu Glu Arg Pro
195 200 205
Glu Leu Phe Gly Ile Asn Pro Glu Thr Gly Asn Glu Gln Met Ser Asn
210 215 220
Leu Asn Phe Thr Gln Gly Ile Ala Gln Ala Leu Trp His Lys Lys Leu
225 230 235 240
Phe His Ile Asp Leu Asn Gly Gln His Gly Pro Lys Phe Asp Gln Asp
245 250 255
Leu Val Phe Gly His Gly Asp Leu Leu Asn Ala Phe Ser Leu Val Asp
260 265 270
Leu Leu Glu Asn Gly Pro Asp Gly Ala Pro Ala Tyr Asp Gly Pro Arg
275 280 285
His Phe Asp Tyr Lys Pro Ser Arg Thr Glu Asp Tyr Asp Gly Val Trp
290 295 300
Glu Ser Ala Lys Ala Asn Ile Arg Met Tyr Leu Leu Leu Lys Glu Arg
305 310 315 320
Ala Lys Ala Phe Arg Ala Asp Pro Glu Val Gln Glu Ala Leu Ala Ala
325 330 335
Ser Lys Val Ala Glu Leu Lys Thr Pro Thr Leu Asn Pro Gly Glu Gly
340 345 350
Tyr Ala Glu Leu Leu Ala Asp Arg Ser Ala Phe Glu Asp Tyr Asp Ala
355 360 365
Asp Ala Val Gly Ala Lys Gly Phe Gly Phe Val Lys Leu Asn Gln Leu
370 375 380
Ala Ile Glu His Leu Leu Gly Ala Arg
385 390

<210> 53
<211> 348
<212> PRT
<213> Bacteriophage T7

<400> 53
Val Asn Ile Lys Thr Asn Pro Phe Lys Ala Val Ser Phe Val Glu Ser
1 5 10 15
Ala Ile Lys Lys Ala Leu Asp Asn Ala Gly Tyr Leu Ile Ala Glu Ile
20 25 30
Lys Tyr Asp Gly Val Arg Gly Asn Ile Cys Val Asp Asn Thr Ala Asn
35 40 45
Ser Tyr Trp Leu Ser Arg Val Ser Lys Thr Ile Pro Ala Leu Glu His
50 55 60
Leu Asn Gly Phe Asp Val Arg Trp Lys Arg Leu Leu Asn Asp Asp Arg
65 70 75 80
Cys Phe Tyr Lys Asp Gly Phe Met Leu Asp Gly Glu Leu Met Val Lys
85 90 95
Gly Val Asp Phe Asn Thr Gly Ser Gly Leu Leu Arg Thr Lys Trp Thr
100 105 110
Asp Thr Lys Asn Gln Glu Phe His Glu Glu Leu Phe Val Glu Pro Ile
115 120 125
Arg Lys Lys Asp Lys Val Pro Phe Lys Leu His Thr Gly His Leu His
130 135 140
Ile Lys Leu Tyr Ala Ile Leu Pro Leu His Ile Val Glu Ser Gly Glu
145 150 155 160
Asp Cys Asp Val Met Thr Leu Leu Met Gln Glu His Val Lys Asn Met
165 170 175
Leu Pro Leu Leu Gln Glu Tyr Phe Pro Glu Ile Glu Trp Gln Ala Ala
180 185 190
Glu Ser Tyr Glu Val Tyr Asp Met Val Glu Leu Gln Gln Leu Tyr Glu
195 200 205
Gln Lys Arg Ala Glu Gly His Glu Gly Leu Ile Val Lys Asp Pro Met
210 215 220
Cys Ile Tyr Lys Arg Gly Lys Lys Ser Gly Trp Trp Lys Met Lys Pro
225 230 235 240
Glu Asn Glu Ala Asp Gly Ile Ile Gln Gly Leu Val Trp Gly Thr Lys
245 250 255
Gly Leu Ala Asn Glu Gly Lys Val Ile Gly Phe Glu Val Leu Leu Glu
260 265 270
Ser Gly Arg Leu Val Asn Ala Thr Asn Ile Ser Arg Ala Leu Met Asp
275 280 285
Glu Phe Thr Glu Thr Val Lys Glu Ala Thr Leu Ser Gln Trp Gly Phe
290 295 300
Phe Ser Pro Tyr Gly Ile Gly Asp Asn Asp Ala Cys Thr Ile Asn Pro
305 310 315 320
Tyr Asp Gly Trp Ala Cys Gln Ile Ser Tyr Met Glu Glu Thr Pro Asp
325 330 335
Gly Ser Leu Arg His Pro Ser Phe Val Met Phe Arg
340 345

<210> 54
<211> 42
<212> DNA
<213> artificial sequence
<220>
<223> binding site for restr1 and restr2
<220>
<221> CDS
<222> (2)..(40)
<223>

<400> 54
g gtg gta tca gca ggc cac tgc tac aag tcc cgc atc cag gt 42
Val Val Ser Ala Gly His Cys Tyr Lys Ser Arg Ile Gln
1 5 10

<210> 55
<211> 13
<212> PRT
<213> artificial sequence
<220>
<223> binding site for restr1 and restr2

<400> 55
Val Val Ser Ala Gly His Cys Tyr Lys Ser Arg Ile Gln
1 5 10

<210> 56
<211> 42
<212> DNA
<213> artificial sequence
<220>
<223> forward primer restr1

<400> 56
ggtggtatcc gcgggccact gctacaagtc ccggatccag gt 42

<210> 57
<211> 42
<212> DNA
<213> artificial sequence
<220>
<223> reverse primer restr2

<400> 57
acctggatcc gggacttgta gcagtggccc gcggatacca cc 42

<210> 58
<211> 50
<212> DNA
<213> artificial sequence
<220>
<223> binding site for restr3 and restr4
<220>
<221> CDS
<222> (3)..(50)
<223>

<400> 58
cc act ggc acg aag tgc ctc atc tct ggc tgg ggc aac act gcg agc 47
Thr Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly Asn Thr Ala Ser
1 5 10 15
tct 50
Ser

<210> 59
<211> 16
<212> PRT
<213> artificial sequence
<220>
<223> binding site for restr3 and restr4

<400> 59

Thr Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly Asn Thr Ala Ser Ser
1 5 10 15

<210> 60
<211> 50
<212> DNA
<213> artificial sequence
<220>
<223> forward primer restr3

<400> 60
ccactggcac gaagtgcctc atctctggct ggggcaacac tgcgagctct 50

<210> 61
<211> 50
<212> DNA
<213> artificial sequence
<220>
<223> reverse primer restr4

<400> 61
agagctagca gtgttgcccc agccagagat gaggcacttg gtaccagtgg 50

<210> 62
<211> 30
<212> DNA
<213> artificial sequence
<220>
<223> primer puc-forward

<400> 62
ggggtacccc accaccatga atccactcct 30

<210> 63
<211> 30
<212> DNA
<213> artificial sequence
<220>
<223> primer puc-reverse

<400> 63
cgggatccgg tatagagact gaagagatac 30

<210> 64
<211> 39
<212> DNA
<213> artificial sequence
<220>
<223> oligox-SDR1f
<220>
<221> misc_feature
<222> (14)..(31)
<223> any nucleotide
<220>
<221> misc_feature
<222> (14)..(31)
<223> any nucleotide or amino acid residue
<220>

<221> CDS
<222> (2)..(37)
<223>

<400> 64
g ggc cac tgc tac nnn nnn nnn nnn nnn nnn aag tcc cg 39
Gly His Cys Tyr Xaa Xaa Xaa Xaa Xaa Xaa Lys Ser
1 5 10

<210> 65
<211> 12
<212> PRT
<213> artificial sequence
<220>
<221> misc_feature
<222> (5)..(5)
<223> The 'Xaa' at location 5 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (6)..(6)
<223> The 'Xaa' at location 6 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (7)..(7)
<223> The 'Xaa' at location 7 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (8)..(8)
<223> The 'Xaa' at location 8 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (9)..(9)
<223> The 'Xaa' at location 9 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (10)..(10)
<223> The 'Xaa' at location 10 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<223> oligox-SDR1f
<220>
<221> misc_feature
<222> (14)..(31)
<223> any nucleotide

<220>
<221> misc_feature
<222> (14)..(31)
<223> any nucleotide or amino acid residue

<400> 65
Gly His Cys Tyr Xaa Xaa Xaa Xaa Xaa Xaa Lys Ser
1 5 10

<210> 66
<211> 45
<212> DNA
<213> artificial sequence
<220>
<223> oligox-SDR1r
<220>
<221> misc_feature
<222> (16)..(33)
<223> any nucleotide

<400> 66
cgcccggtga cgatgnnnnn nnnnnnnnnn nnnttcaggg cctag 45

<210> 67
<211> 47
<212> DNA
<213> artificial sequence
<220>
<223> oligox-SDR2f
<220>
<221> CDS
<222> (2)..(46)
<223>
<220>
<221> misc_feature
<222> (29)..(43)
<223> any nucleotide or amino acid residue

<400> 67
c aag tgc ctc atc tct ggc tgg ggc aac nnn nnn nnn nnn nnn act g 47
Lys Cys Leu Ile Ser Gly Trp Gly Asn Xaa Xaa Xaa Xaa Xaa Thr
1 5 10 15

<210> 68
<211> 15
<212> PRT
<213> artificial sequence
<220>
<221> misc_feature
<222> (10)..(10)
<223> The 'Xaa' at location 10 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (11)..(11)
<223> The 'Xaa' at location 11 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (12)..(12)
<223> The 'Xaa' at location 12 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (13)..(13)
<223> The 'Xaa' at location 13 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<221> misc_feature
<222> (14)..(14)
<223> The 'Xaa' at location 14 stands for Lys, Asn, Arg, Ser, Thr, Ile, Met, Glu, Asp, Gly, Ala, Val, Gln, His, Pro, Leu, a stop codon, Tyr, Trp, Cys, or Phe.

<220>
<223> oligox-SDR2f
<220>
<221> misc_feature
<222> (29)..(43)
<223> any nucleotide or amino acid residue

<400> 68
Lys Cys Leu Ile Ser Gly Trp Gly Asn Xaa Xaa Xaa Xaa Xaa Thr
1 5 10 15

<210> 69
<211> 55
<212> DNA
<213> artificial sequence
<220>
<223> oligox-SDR2r
<220>
<221> misc_feature
<222> (33)..(47)
<223> any base
<220>
<221> misc_feature
<222> (33)..(47)
<223> any nucleotide

<400> 69
catggttcac ggagtagaga ccgaccccgt tgnnnnnnnn nnnnnnntga cgatc 55

<210> 70
<211> 59
<212> DNA
<213> artificial sequence
<220>
<223> primer SDR1-mutnnb-forward
<220>
<221> misc_feature
<222> (24)..(40)
<223> N=A, C, G, T; B=C, G, T; V=A, C, G

<400> 70
tggtatccgc gggccactgc tacnnbnnbn nbnnbnnbnn baagtcccgg atccaggtg 59

<210> 71
<211> 52
<212> DNA
<213> artificial sequence
<220>
<223> primer SDR2-mutnnb-reverse
<220>
<221> misc_feature
<222> (20)..(33)
<223> N=A, C, G, T; B=C, G, T; V=A, C, G

<400> 71
ggcgccagag ctagcagtvn nvnnvnnvnn vnngttgccc cagccagaga tg 52

<210> 72
<211> 6
<212> PRT
<213> artificial sequence

<220>
<223> variant g SDR1
<400> 72
Ala Phe Phe Asn Gly Asp
1 5

<210> 73
<211> 5
<212> PRT
<213> artificial sequence

<220>
<223> variant g SDR2
<400> 73
Arg Lys Asp Pro Trp
1 5
<210> 74
<211> 234
<212> PRT
<213> artificial sequence

<220>
<223> artificial sequence

<400> 74

Ile Val Gly Gly Tyr Asn Cys Glu Glu Asn Ser Val Pro Tyr Gln Val
1 5 10 15
Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Glu
20 25 30
Gln Trp Val Val Ser Ala Gly His Cys Tyr Ala Ala Phe Asn Gly Lys
35 40 45
Ser Arg Ile Gln Val Arg Leu Gly Glu His Asn Ile Glu Val Leu Glu
50 55 60
Gly Asn Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro Gln
65 70 75 80
Tyr Asp Arg Lys Thr Leu Asn Asn Asp Ile Met Leu Ile Lys Leu Ser
85 90 95
Ser Arg Ala Val Ile Asn Ala Arg Val Ser Thr Ile Ser Leu Pro Thr
100 105 110
Ala Pro Pro Ala Thr Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly Asn
115 120 125
Arg Lys Asp Phe Trp Thr Ala Ser Ser Gly Ala Asp Tyr Pro Asp Glu
130 135 140
Leu Gln Cys Leu Asp Ala Pro Val Leu Ser Gln Ala Lys Cys Glu Ala
145 150 155 160
Ser Tyr Pro Gly Lys Ile Thr Ser Asn Met Phe Cys Val Gly Phe Leu
165 170 175
Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Val Val
180 185 190
Cys Asn Gly Gln Leu Gln Gly Val Val Ser Trp Gly Asp Gly Cys Ala
195 200 205
Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys Val Tyr Asn Tyr Val Lys
210 215 220
Trp Ile Lys Asn Thr Ile Ala Ala Asn Ser
225 230

<210> 75
<211> 234
<212> PRT
<213> artificial sequence

<220>
<223> artificial sequence
<400> 75
Ile Val Gly Gly Tyr Asn Cys Glu Glu Asn Ser Val Pro Tyr Gln Val

```
1               5                   10                  15
Ser Leu Asn Ser Gly Tyr His Phe Cys Gly Gly Ser Leu Ile Asn Glu
            20                  25                  30
Gln Trp Val Val Ser Ala Gly His Cys Tyr Ala Ala Phe Asn Gly Lys
        35                  40                  45
Ser Arg Ile Gln Val Arg Leu Gly Glu His Asn Ile Gly Val Leu Glu
    50                  55                  60
Gly Asn Glu Gln Phe Ile Asn Ala Ala Lys Ile Ile Arg His Pro Gln
65                  70                  75                  80
Tyr Asp Trp Lys Thr Leu Asn Asn Asp Ile Met Leu Ile Lys Leu Ser
                85                  90                  95
Ser Arg Ala Val Ile Asn Ala Arg Val Ser Thr Ile Ser Leu Pro Thr
            100                 105                 110
Ala Pro Pro Ala Thr Gly Thr Lys Cys Leu Ile Ser Gly Trp Gly Asn
        115                 120                 125
Arg Lys Asp Phe Trp Thr Ala Ser Ser Gly Ala Asp Phe Pro Asp Glu
    130                 135                 140
Leu Gln Cys Leu Asp Ala Pro Val Leu Ser Gln Thr Lys Cys Glu Ala
145                 150                 155                 160
Ser Tyr Pro Gly Lys Ile Thr Ser Asn Met Phe Cys Val Gly Phe Leu
                165                 170                 175
Glu Gly Gly Lys Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Val Val
            180                 185                 190
Arg Asn Gly Gln Leu Gln Gly Val Val Ser Trp Gly Asp Gly Cys Ala
        195                 200                 205
Gln Lys Asn Lys Pro Gly Val Tyr Thr Lys Val Tyr Asn Tyr Val Lys
    210                 215                 220
Trp Ile Lys Asn Thr Ile Ala Ala Asn Ser
225                 230
```

<210> 76
<211> 12
<212> PRT
<213> artificial sequence
<220>
<223> substrate A

<400> 76
```
Leu Leu Trp Leu Gly Arg Val Val Gly Gly Pro Val
1               5                   10
```

<210> 77
<211> 12
<212> PRT
<213> artificial sequence
<220>

<223> substrate B

<400> 77
Lys Lys Trp Leu Gly Arg Val Pro Gly Gly Pro Val
1 5 10

<210> 78
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant1 SDR1

<400> 78
Asp Ala Val Gly Arg Asp
1 5

<210> 79
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant2 SDR1

<400> 79
Asn Gly Arg Asp Leu Glu
1 5

<210> 80
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant3 SDR1

<400> 80
Gly Phe Val Met Phe Asn
1 5

<210> 81
<211> 5
<212> PRT
<213> artificial sequence
<220>
<223> variant1 SDR2

<400> 81
Arg Val His Pro Ser
1 5

<210> 82
<211> 5
<212> PRT
<213> artificial sequence
<220>
<223> variant2 SDR2

<400> 82
Val Arg Gly Thr Trp
1 5

<210> 83
<211> 5
<212> PRT
<213> artificial sequence
<220>

<223> variant3 SDR2

<400> 83
Arg Ser Pro Leu Thr
1 5

<210> 84
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant a SDR1

<400> 84
Arg Pro Trp Asp Pro Ser
1 5

<210> 85
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant b SDR1

<400> 85
Gly Phe Val Met Phe Asn
1 5

<210> 86
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant c SDR1

<400> 86
Glu Ile Ala Asn Arg Glu
1 5

<210> 87
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant d SDR1

<400> 87
Lys Ala Val Val Gly Thr

1 5

<210> 88
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant e SDR1

<400> 88
Val Asn Ile Met Ala Ala
1 5

<210> 89
<211> 6
<212> PRT
<213> artificial sequence
<220>
<223> variant f SDR1

<400> 89
Ala Ala Phe Asn Gly Asp
1 5

<210> 90
<211> 5
<212> PRT
<213> artificial sequence
<220>
<223> variant a SDR2
<400> 90

Val His Pro Thr Ser
1 5

<210> 91
<211> 5
<212> PRT
<213> artificial sequence
<220>
<223> variant b SDR2

<400> 91
Arg Ser Pro Leu Thr
1 5

<210> 92
<211> 5
<212> PRT
<213> artificial sequence
<220>
<223> variant c SDR2

<400> 92
Arg Gly Ala Arg Thr
1 5

<210> 93
<211> 5
<212> PRT
<213> artificial sequence
<220>
<223> variant d SDR2

<400> 93
Arg Thr Pro Ile Ser
1 5

<210> 94
<211> 5
<212> PRT
<213> artificial sequence
<220>
<223> variant e SDR2

<400> 94
Thr Thr Ala Arg Lys
1 5

<210> 95
<211> 5
<212> PRT
<213> artificial sequence
<220>
<223> variant f SDR2

<400> 95
Arg Lys Asp Phe Trp
1 5

<210> 96
<211> 157
<212> PRT
<213> Homo sapiens

<400> 96
Val Arg Ser Ser Ser Arg Thr Pro Ser Asp Lys Pro Val Ala His Val
1 5 10 15
Val Ala Asn Pro Gln Ala Glu Gly Gln Leu Gln Trp Leu Asn Arg Arg
20 25 30
Ala Asn Ala Leu Leu Ala Asn Gly Val Glu Leu Arg Asp Asn Gln Leu
35 40 45
Val Val Pro Ser Glu Gly Leu Tyr Leu Ile Tyr Ser Gln Val Leu Phe
50 55 60
Lys Gly Gln Gly Cys Pro Ser Thr His Val Leu Leu Thr His Thr Ile
65 70 75 80
Ser Arg Ile Ala Val Ser Tyr Gln Thr Lys Val Asn Leu Leu Ser Ala
85 90 95
Ile Lys Ser Pro Cys Gln Arg Glu Thr Pro Glu Gly Ala Glu Ala Lys
100 105 110
Pro Trp Tyr Glu Pro Ile Tyr Leu Gly Gly Val Phe Gln Leu Glu Lys
115 120 125
Gly Asp Arg Leu Ser Ala Glu Ile Asn Arg Pro Asp Tyr Leu Leu Phe
130 135 140
Ala Glu Ser Gly Gln Val Tyr Phe Gly Ile Ile Ala Leu
145 150 155

<210> 97
<211> 306
<212> PRT
<213> Homo sapiens

<400> 97
Ile Trp Glu Leu Lys Lys Asp Val Tyr Val Val Glu Leu Asp Trp Tyr
1 5 10 15
Pro Asp Ala Pro Gly Glu Met Val Val Leu Thr Cys Asp Thr Pro Glu
20 25 30
Glu Asp Gly Ile Thr Trp Thr Leu Asp Gln Ser Ser Glu Val Leu Gly
35 40 45
Ser Gly Lys Thr Leu Thr Ile Gln Val Lys Glu Phe Gly Asp Ala Gly
50 55 60
Gln Tyr Thr Cys His Lys Gly Gly Glu Val Leu Ser His Ser Leu Leu
65 70 75 80
Leu Leu His Lys Lys Glu Asp Gly Ile Trp Ser Thr Asp Ile Leu Lys
85 90 95
Asp Gln Lys Glu Pro Lys Asn Lys Thr Phe Leu Arg Cys Glu Ala Lys
100 105 110
Asn Tyr Ser Gly Arg Phe Thr Cys Trp Trp Leu Thr Thr Ile Ser Thr
115 120 125
Asp Leu Thr Phe Ser Val Lys Ser Ser Arg Gly Ser Ser Asp Pro Gln
130 135 140
Gly Val Thr Cys Gly Ala Ala Thr Leu Ser Ala Glu Arg Val Arg Gly
145 150 155 160
Asp Asn Lys Glu Tyr Glu Tyr Ser Val Glu Cys Gln Glu Asp Ser Ala
165 170 175
Cys Pro Ala Ala Glu Glu Ser Leu Pro Ile Glu Val Met Val Asp Ala
180 185 190
Val His Lys Leu Lys Tyr Glu Asn Tyr Thr Ser Ser Phe Phe Ile Arg
195 200 205
Asp Ile Ile Lys Pro Asp Pro Pro Lys Asn Leu Gln Leu Lys Pro Leu
210 215 220
Lys Asn Ser Arg Gln Val Glu Val Ser Trp Glu Tyr Pro Asp Thr Trp
225 230 235 240
Ser Thr Pro His Ser Tyr Phe Ser Leu Thr Phe Cys Val Gln Val Gln
245 250 255
Gly Lys Ser Lys Arg Glu Lys Lys Asp Arg Val Phe Thr Asp Lys Thr
260 265 270
Ser Ala Thr Val Ile Cys Arg Lys Asn Ala Ser Ile Ser Val Arg Ala
275 280 285
Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser Val Pro
290 295 300
Cys Ser
305

<210> 98
<211> 157
<212> PRT
<213> Homo sapiens

<400> 98
Tyr Phe Gly Lys Leu Glu Ser Lys Leu Ser Val Ile Arg Asn Leu Asn
1 5 10 15
Asp Gln Val Leu Phe Ile Asp Gln Gly Asn Arg Pro Leu Phe Glu Asp
20 25 30
Met Thr Asp Ser Asp Cys Arg Asp Asn Ala Pro Arg Thr Ile Phe Ile
35 40 45
Ile Ser Met Tyr Lys Asp Ser Gln Pro Arg Gly Met Ala Val Thr Ile
50 55 60

```
Ser Val Lys Cys Glu Lys Ile Ser Thr Leu Ser Cys Glu Asn Lys Ile
65                  70                  75                  80
Ile Ser Phe Lys Glu Met Asn Pro Pro Asp Asn Ile Lys Asp Thr Lys
                85                  90                  95
Ser Asp Ile Ile Phe Phe Gln Arg Ser Val Pro Gly His Asp Asn Lys
            100                 105                 110
Met Gln Phe Glu Ser Ser Ser Tyr Glu Gly Tyr Phe Leu Ala Cys Glu
        115                 120                 125
Lys Glu Arg Asp Leu Phe Lys Leu Ile Leu Lys Lys Glu Asp Glu Leu
    130                 135                 140
Gly Asp Arg Ser Ile Met Phe Thr Val Gln Asn Glu Asp
145                 150                 155


<210>  99
<211>  133
<212>  PRT
<213>  Homo sapiens

<400>  99
Ala Pro Thr Ser Ser Ser Thr Lys Lys Thr Gln Leu Gln Leu Glu His
1               5                   10                  15
Leu Leu Leu Asp Leu Gln Met Ile Leu Asn Gly Ile Asn Asn Tyr Lys
            20                  25                  30
Asn Pro Lys Leu Thr Arg Met Leu Thr Phe Lys Phe Tyr Met Pro Lys
        35                  40                  45
Lys Ala Thr Glu Leu Lys His Leu Gln Cys Leu Glu Glu Glu Leu Lys
    50                  55                  60
Pro Leu Glu Glu Val Leu Asn Leu Ala Gln Ser Lys Asn Phe His Leu
65                  70                  75                  80
Arg Pro Arg Asp Leu Ile Ser Asn Ile Asn Val Ile Val Leu Glu Leu
                85                  90                  95
Lys Gly Ser Glu Thr Thr Phe Met Cys Glu Tyr Ala Asp Glu Thr Ala
            100                 105                 110
Thr Ile Val Glu Phe Leu Asn Arg Trp Ile Thr Phe Cys Gln Ser Ile
        115                 120                 125
Ile Ser Thr Leu Thr
    130


<210>  100
<211>  72
<212>  PRT
<213>  Homo sapiens

<400>  100
Ser Ala Lys Glu Leu Arg Cys Gln Cys Ile Lys Thr Tyr Ser Lys Pro
1               5                   10                  15
Phe His Pro Lys Phe Ile Lys Glu Leu Arg Val Ile Glu Ser Gly Pro
            20                  25                  30
His Cys Ala Asn Thr Glu Ile Ile Val Lys Leu Ser Asp Gly Arg Glu
        35                  40                  45
Leu Cys Leu Asp Pro Lys Glu Asn Trp Val Gln Arg Val Val Glu Lys
    50                  55                  60
Phe Leu Lys Arg Ala Glu Asn Ser
65                  70


<210>  101
<211>  74
<212>  PRT
<213>  Homo sapiens
```

<400> 101
Gly Pro Ala Ser Val Pro Thr Thr Cys Cys Phe Asn Leu Ala Asn Arg
1 5 10 15
Lys Ile Pro Leu Gln Arg Leu Glu Ser Tyr Arg Arg Ile Thr Ser Gly
20 25 30
Lys Cys Pro Gln Lys Ala Val Ile Phe Lys Thr Lys Leu Ala Lys Asp
35 40 45
Ile Cys Ala Asp Pro Lys Lys Lys Trp Val Gln Asp Ser Met Lys Tyr
50 55 60
Leu Asp Gln Lys Ser Pro Thr Pro Lys Pro
65 70

<210> 102
<211> 76
<212> PRT
<213> Homo sapiens

<400> 102
Gln Pro Asp Ala Ile Asn Ala Pro Val Thr Cys Cys Tyr Asn Phe Thr
1 5 10 15
Asn Arg Lys Ile Ser Val Gln Arg Leu Ala Ser Tyr Arg Arg Ile Thr
20 25 30
Ser Ser Lys Cys Pro Lys Glu Ala Val Ile Phe Lys Thr Ile Val Ala
35 40 45
Lys Glu Ile Cys Ala Asp Pro Lys Gln Lys Trp Val Gln Asp Ser Met
50 55 60
Asp His Leu Asp Lys Gln Thr Gln Thr Pro Lys Thr
65 70 75

<210> 103
<211> 206
<212> PRT
<213> Homo sapiens

<400> 103
Ala Pro Met Ala Glu Gly Gly Gly Gln Asn His His Glu Val Val Lys
1 5 10 15
Phe Met Asp Val Tyr Gln Arg Ser Tyr Cys His Pro Ile Glu Thr Leu
20 25 30
Val Asp Ile Phe Gln Glu Tyr Pro Asp Glu Ile Glu Tyr Ile Phe Lys
35 40 45
Pro Ser Cys Val Pro Leu Met Arg Cys Gly Gly Cys Cys Asn Asp Glu
50 55 60
Gly Leu Glu Cys Val Pro Thr Glu Glu Ser Asn Ile Thr Met Gln Ile
65 70 75 80
Met Arg Ile Lys Pro His Gln Gly Gln His Ile Gly Glu Met Ser Phe
85 90 95
Leu Gln His Asn Lys Cys Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg
100 105 110
Gln Glu Lys Lys Ser Val Arg Gly Lys Gly Lys Gly Gln Lys Arg Lys
115 120 125
Arg Lys Lys Ser Arg Tyr Lys Ser Trp Ser Val Tyr Val Gly Ala Arg
130 135 140
Cys Cys Leu Met Pro Trp Ser Leu Pro Gly Pro His Pro Cys Gly Pro
145 150 155 160
Cys Ser Glu Arg Arg Lys His Leu Phe Val Gln Asp Pro Gln Thr Cys
165 170 175
Lys Cys Ser Cys Lys Asn Thr Asp Ser Arg Cys Lys Ala Arg Gln Leu
180 185 190
Glu Leu Asn Glu Arg Thr Cys Arg Cys Asp Lys Pro Arg Arg
195 200 205

<210> 104
<211> 112
<212> PRT
<213> Homo sapiens

<400> 104

Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys Asn Cys Cys
1 5 10 15
Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly Trp Lys Trp
20 25 30
Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu Gly Pro Cys
35 40 45
Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val Leu Ala Leu
50 55 60
Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys Cys Val Pro
65 70 75 80
Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly Arg Lys Pro
85 90 95
Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys Lys Cys Ser
100 105 110

<210> 105
<211> 30
<212> PRT
<213> Homo sapiens

<400> 105
Phe Val Asn Gln His Leu Cys Gly Ser His Leu Val Glu Ala Leu Tyr
1 5 10 15
Leu Val Cys Gly Glu Arg Gly Phe Phe Tyr Thr Pro Lys Thr
20 25 30

<210> 106
<211> 21
<212> PRT
<213> Homo sapiens

<400> 106
Gly Ile Val Glu Gln Cys Cys Thr Ser Ile Cys Ser Leu Tyr Gln Leu
1 5 10 15
Glu Asn Tyr Cys Asn
20

<210> 107
<211> 28
<212> PRT
<213> Homo sapiens

<400> 107
Gly Ser Ser Phe Leu Ser Pro Glu His Gln Arg Val Gln Gln Arg Lys
1 5 10 15
Glu Ser Lys Lys Pro Pro Ala Lys Leu Gln Pro Arg
20 25

<210> 108
<211> 9

<212> PRT
<213> Homo sapiens

<400> 108
Arg Val Tyr Ile His Pro Phe His Leu
1 5

<210> 109
<211> 114
<212> PRT
<213> Homo sapiens

<400> 109
Pro Met Phe Ile Val Asn Thr Asn Val Pro Arg Ala Ser Val Pro Asp
1 5 10 15
Gly Phe Leu Ser Glu Leu Thr Gln Gln Leu Ala Gln Ala Thr Gly Lys
20 25 30
Pro Pro Gln Tyr Ile Ala Val His Val Val Pro Asp Gln Leu Met Ala
35 40 45
Phe Gly Gly Ser Ser Glu Pro Cys Ala Leu Cys Ser Leu His Ser Ile
50 55 60
Gly Lys Ile Gly Gly Ala Gln Asn Arg Ser Tyr Ser Lys Leu Leu Cys
65 70 75 80
Gly Leu Leu Ala Glu Arg Leu Arg Ile Ser Pro Asp Arg Val Tyr Ile
85 90 95
Asn Tyr Tyr Asp Met Asn Ala Ala Asn Val Gly Trp Asn Asn Ser Thr
100 105 110
Phe Ala

<210> 110
<211> 425
<212> PRT
<213> Homo sapiens

<400> 110
Met Gly Pro Arg Arg Leu Leu Leu Val Ala Ala Cys Phe Ser Leu Cys
1 5 10 15
Gly Pro Leu Leu Ser Ala Arg Thr Arg Ala Arg Arg Pro Glu Ser Lys
20 25 30
Ala Thr Asn Ala Thr Leu Asp Pro Arg Ser Phe Leu Leu Arg Asn Pro
35 40 45
Asn Asp Lys Tyr Glu Pro Phe Trp Glu Asp Glu Glu Lys Asn Glu Ser
50 55 60
Gly Leu Thr Glu Tyr Arg Leu Val Ser Ile Asn Lys Ser Ser Pro Leu
65 70 75 80
Gln Lys Gln Leu Pro Ala Phe Ile Ser Glu Asp Ala Ser Gly Tyr Leu
85 90 95
Thr Ser Ser Trp Leu Thr Leu Phe Val Pro Ser Val Tyr Thr Gly Val
100 105 110
Phe Val Val Ser Leu Pro Leu Asn Ile Met Ala Ile Val Val Phe Ile
115 120 125
Leu Lys Met Lys Val Lys Lys Pro Ala Val Val Tyr Met Leu His Leu
130 135 140
Ala Thr Ala Asp Val Leu Phe Val Ser Val Leu Pro Phe Lys Ile Ser
145 150 155 160
Tyr Tyr Phe Ser Gly Ser Asp Trp Gln Phe Gly Ser Glu Leu Cys Arg
165 170 175
Phe Val Thr Ala Ala Phe Tyr Cys Asn Met Tyr Ala Ser Ile Leu Leu
180 185 190
Met Thr Val Ile Ser Ile Asp Arg Phe Leu Ala Val Val Tyr Pro Met

```
        195                 200                 205
Gln Ser Leu Ser Trp Arg Thr Leu Gly Arg Ala Ser Phe Thr Cys Leu
    210                 215                 220
Ala Ile Trp Ala Leu Ala Ile Ala Gly Val Val Pro Leu Leu Leu Lys
225                 230                 235                 240
Glu Gln Thr Ile Gln Val Pro Gly Leu Asn Ile Thr Thr Cys His Asp
                245                 250                 255
Val Leu Asn Glu Thr Leu Leu Glu Gly Tyr Tyr Ala Tyr Tyr Phe Ser
            260                 265                 270
Ala Phe Ser Ala Val Phe Phe Phe Val Pro Leu Ile Ile Ser Thr Val
        275                 280                 285
Cys Tyr Val Ser Ile Ile Arg Cys Leu Ser Ser Ser Ala Val Ala Asn
    290                 295                 300
Arg Ser Lys Lys Ser Arg Ala Leu Phe Leu Ser Ala Ala Val Phe Cys
305                 310                 315                 320
Ile Phe Ile Ile Cys Phe Gly Pro Thr Asn Val Leu Leu Ile Ala His
                325                 330                 335
Tyr Ser Phe Leu Ser His Thr Ser Thr Thr Glu Ala Ala Tyr Phe Ala
            340                 345                 350
Tyr Leu Leu Cys Val Cys Val Ser Ser Ile Ser Cys Cys Ile Asp Pro
        355                 360                 365
Leu Ile Tyr Tyr Tyr Ala Ser Ser Glu Cys Gln Arg Tyr Val Tyr Ser
    370                 375                 380
Ile Leu Cys Cys Lys Glu Ser Ser Asp Pro Ser Ser Tyr Asn Ser Ser
385                 390                 395                 400
Gly Gln Leu Met Ala Ser Lys Met Asp Thr Cys Ser Ser Asn Leu Asn
                405                 410                 415
Asn Ser Ile Tyr Lys Lys Leu Leu Thr
            420                 425
```

<210> 111
<211> 397
<212> PRT
<213> Homo sapiens

<400> 111

```
Met Arg Ser Pro Ser Ala Ala Trp Leu Leu Gly Ala Ala Ile Leu Leu
1               5                   10                  15
Ala Ala Ser Leu Ser Cys Ser Gly Thr Ile Gln Gly Thr Asn Arg Ser
            20                  25                  30
Ser Lys Gly Arg Ser Leu Ile Gly Lys Val Asp Gly Thr Ser His Val
        35                  40                  45
Thr Gly Lys Gly Val Thr Val Glu Thr Val Phe Ser Val Asp Glu Phe
    50                  55                  60
Ser Ala Ser Val Leu Thr Gly Lys Leu Thr Thr Val Phe Leu Pro Ile
65                  70                  75                  80
Val Tyr Thr Ile Val Phe Val Val Gly Leu Pro Ser Asn Gly Met Ala
                85                  90                  95
Leu Trp Val Phe Leu Phe Arg Thr Lys Lys Lys His Pro Ala Val Ile
            100                 105                 110
Tyr Met Ala Asn Leu Ala Leu Ala Asp Leu Leu Ser Val Ile Trp Phe
        115                 120                 125
Pro Leu Lys Ile Ala Tyr His Ile His Gly Asn Asn Trp Ile Tyr Gly
    130                 135                 140
Glu Ala Leu Cys Asn Val Leu Ile Gly Phe Phe Tyr Gly Asn Met Tyr
145                 150                 155                 160
Cys Ser Ile Leu Phe Met Thr Cys Leu Ser Val Gln Arg Tyr Trp Val
                165                 170                 175
Ile Val Asn Pro Met Gly His Ser Arg Lys Lys Ala Asn Ile Ala Ile
            180                 185                 190
Gly Ile Ser Leu Ala Ile Trp Leu Leu Ile Leu Leu Val Thr Ile Pro
        195                 200                 205
```

```
Leu Tyr Val Val Lys Gln Thr Ile Phe Ile Pro Ala Leu Asn Ile Thr
    210                 215                 220
Thr Cys His Asp Val Leu Pro Glu Gln Leu Leu Val Gly Asp Met Phe
225                 230                 235                 240
Asn Tyr Phe Leu Ser Leu Ala Ile Gly Val Phe Leu Phe Pro Ala Phe
                245                 250                 255
Leu Thr Ala Ser Ala Tyr Val Leu Met Ile Arg Met Leu Arg Ser Ser
            260                 265                 270
Ala Met Asp Glu Asn Ser Glu Lys Lys Arg Lys Arg Ala Ile Lys Leu
        275                 280                 285
Ile Val Thr Val Leu Ala Met Tyr Leu Ile Cys Phe Thr Pro Ser Asn
    290                 295                 300
Leu Leu Leu Val Val His Tyr Phe Leu Ile Lys Ser Gln Gly Gln Ser
305                 310                 315                 320
His Val Tyr Ala Leu Tyr Ile Val Ala Leu Cys Leu Ser Thr Leu Asn
                325                 330                 335
Ser Cys Ile Asp Pro Phe Val Tyr Tyr Phe Val Ser His Asp Phe Arg
            340                 345                 350
Asp His Ala Lys Asn Ala Leu Leu Cys Arg Ser Val Arg Thr Val Lys
        355                 360                 365
Gln Met Gln Val Ser Leu Thr Ser Lys Lys His Ser Arg Lys Ser Ser
    370                 375                 380
Ser Tyr Ser Ser Ser Ser Thr Thr Val Lys Thr Ser Tyr
385                 390                 395
```

<210> 112
<211> 153
<212> PRT
<213> Homo sapiens

<400> 112

```
Ala Pro Val Arg Ser Leu Asn Cys Thr Leu Arg Asp Ser Gln Gln Lys
1               5                   10                  15
Ser Leu Val Met Ser Gly Pro Tyr Glu Leu Lys Ala Leu His Leu Gln
            20                  25                  30
Gly Gln Asp Met Glu Gln Gln Val Val Phe Ser Met Ser Phe Val Gln
        35                  40                  45
Gly Glu Glu Ser Asn Asp Lys Ile Pro Val Ala Leu Gly Leu Lys Glu
    50                  55                  60
Lys Asn Leu Tyr Leu Ser Cys Val Leu Lys Asp Asp Lys Pro Thr Leu
65                  70                  75                  80
Gln Leu Glu Ser Val Asp Pro Lys Asn Tyr Pro Lys Lys Lys Met Glu
                85                  90                  95
Lys Arg Phe Val Phe Asn Lys Ile Glu Ile Asn Asn Lys Leu Glu Phe
            100                 105                 110
Glu Ser Ala Gln Phe Pro Asn Trp Tyr Ile Ser Thr Ser Gln Ala Glu
        115                 120                 125
Asn Met Pro Val Phe Leu Gly Gly Thr Lys Gly Gly Gln Asp Ile Thr
    130                 135                 140
Asp Phe Thr Met Gln Phe Val Ser Ser
145                 150
```

<210> 113
<211> 385
<212> PRT
<213> Homo sapiens

<400> 113

```
Met Trp Gly Arg Leu Leu Leu Trp Pro Leu Val Leu Gly Phe Ser Leu
1               5                   10                  15
```

```
Ser Gly Gly Thr Gln Thr Pro Ser Val Tyr Asp Glu Ser Gly Ser Thr
            20                  25                  30
Gly Gly Gly Asp Asp Ser Thr Pro Ser Ile Leu Pro Ala Pro Arg Gly
        35                  40                  45
Tyr Pro Gly Gln Val Cys Ala Asn Asp Ser Asp Thr Leu Glu Leu Pro
    50                  55                  60
Asp Ser Ser Arg Ala Leu Leu Leu Gly Trp Val Pro Thr Arg Leu Val
65                  70                  75                  80
Pro Ala Leu Tyr Gly Leu Val Leu Val Val Gly Leu Pro Ala Asn Gly
                85                  90                  95
Leu Ala Leu Trp Val Leu Ala Thr Gln Ala Pro Arg Leu Pro Ser Thr
            100                 105                 110
Met Leu Leu Met Asn Leu Ala Thr Ala Asp Leu Leu Leu Ala Leu Ala
        115                 120                 125
Leu Pro Pro Arg Ile Ala Tyr His Leu Arg Gly Gln Arg Trp Pro Phe
    130                 135                 140
Gly Glu Ala Ala Cys Arg Leu Ala Thr Ala Ala Leu Tyr Gly His Met
145                 150                 155                 160
Tyr Gly Ser Val Leu Leu Leu Ala Ala Val Ser Leu Asp Arg Tyr Leu
                165                 170                 175
Ala Leu Val His Pro Leu Arg Ala Arg Ala Leu Arg Gly Arg Arg Leu
            180                 185                 190
Ala Leu Gly Leu Cys Met Ala Ala Trp Leu Met Ala Ala Ala Leu Ala
        195                 200                 205
Leu Pro Leu Thr Leu Gln Arg Gln Thr Phe Arg Leu Ala Arg Ser Asp
    210                 215                 220
Arg Val Leu Cys His Asp Ala Leu Pro Leu Asp Ala Gln Ala Ser His
225                 230                 235                 240
Trp Gln Pro Ala Phe Thr Cys Leu Ala Leu Leu Gly Cys Phe Leu Pro
                245                 250                 255
Leu Leu Ala Met Leu Leu Cys Tyr Gly Ala Thr Leu His Thr Leu Ala
            260                 265                 270
Ala Ser Gly Arg Arg Tyr Gly His Ala Leu Arg Leu Thr Ala Val Val
        275                 280                 285
Leu Ala Ser Ala Val Ala Phe Phe Val Pro Ser Asn Leu Leu Leu Leu
    290                 295                 300
Leu His Tyr Ser Asp Pro Ser Pro Ser Ala Trp Gly Asn Leu Tyr Gly
305                 310                 315                 320
Ala Tyr Val Pro Ser Leu Ala Leu Ser Thr Leu Asn Ser Cys Val Asp
                325                 330                 335
Pro Phe Ile Tyr Tyr Tyr Val Ser Ala Glu Phe Arg Asp Lys Val Arg
            340                 345                 350
Ala Gly Leu Phe Gln Arg Ser Pro Gly Asp Thr Val Ala Ser Lys Ala
        355                 360                 365
Ser Ala Glu Gly Gly Ser Arg Gly Met Gly Thr His Ser Ser Leu Leu
    370                 375                 380
Gln
385
```

<210> 114
<211> 1338
<212> PRT
<213> Homo sapiens

<400> 114

```
Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser
1               5                   10                  15
Cys Leu Leu Leu Thr Gly Ser Ser Ser Gly Ser Lys Leu Lys Asp Pro
            20                  25                  30
Glu Leu Ser Leu Lys Gly Thr Gln His Ile Met Gln Ala Gly Gln Thr
        35                  40                  45
Leu His Leu Gln Cys Arg Gly Glu Ala Ala His Lys Trp Ser Leu Pro
```

```
    50                  55                  60
Glu Met Val Ser Lys Glu Ser Glu Arg Leu Ser Ile Thr Lys Ser Ala
65                  70                  75                  80
Cys Gly Arg Asn Gly Lys Gln Phe Cys Ser Thr Leu Thr Leu Asn Thr
                85                  90                  95
Ala Gln Ala Asn His Thr Gly Phe Tyr Ser Cys Lys Tyr Leu Ala Val
            100                 105                 110
Pro Thr Ser Lys Lys Lys Glu Thr Glu Ser Ala Ile Tyr Ile Phe Ile
        115                 120                 125
Ser Asp Thr Gly Arg Pro Phe Val Glu Met Tyr Ser Glu Ile Pro Glu
    130                 135                 140
Ile Ile His Met Thr Glu Gly Arg Glu Leu Val Ile Pro Cys Arg Val
145                 150                 155                 160
Thr Ser Pro Asn Ile Thr Val Thr Leu Lys Lys Phe Pro Leu Asp Thr
                165                 170                 175
Leu Ile Pro Asp Gly Lys Arg Ile Ile Trp Asp Ser Arg Lys Gly Phe
            180                 185                 190
Ile Ile Ser Asn Ala Thr Tyr Lys Glu Ile Gly Leu Leu Thr Cys Glu
        195                 200                 205
Ala Thr Val Asn Gly His Leu Tyr Lys Thr Asn Tyr Leu Thr His Arg
    210                 215                 220
Gln Thr Asn Thr Ile Ile Asp Val Gln Ile Ser Thr Pro Arg Pro Val
225                 230                 235                 240
Lys Leu Leu Arg Gly His Thr Leu Val Leu Asn Cys Thr Ala Thr Thr
                245                 250                 255
Pro Leu Asn Thr Arg Val Gln Met Thr Trp Ser Tyr Pro Asp Glu Lys
            260                 265                 270
Asn Lys Arg Ala Ser Val Arg Arg Arg Ile Asp Gln Ser Asn Ser His
        275                 280                 285
Ala Asn Ile Phe Tyr Ser Val Leu Thr Ile Asp Lys Met Gln Asn Lys
    290                 295                 300
Asp Lys Gly Leu Tyr Thr Cys Arg Val Arg Ser Gly Pro Ser Phe Lys
305                 310                 315                 320
Ser Val Asn Thr Ser Val His Ile Tyr Asp Lys Ala Phe Ile Thr Val
                325                 330                 335
Lys His Arg Lys Gln Gln Val Leu Glu Thr Val Ala Gly Lys Arg Ser
            340                 345                 350
Tyr Arg Leu Ser Met Lys Val Lys Ala Phe Pro Ser Pro Glu Val Val
        355                 360                 365
Trp Leu Lys Asp Gly Leu Pro Ala Thr Glu Lys Ser Ala Arg Tyr Leu
    370                 375                 380
Thr Arg Gly Tyr Ser Leu Ile Ile Lys Asp Val Thr Glu Glu Asp Ala
385                 390                 395                 400
Gly Asn Tyr Thr Ile Leu Leu Ser Ile Lys Gln Ser Asn Val Phe Lys
                405                 410                 415
Asn Leu Thr Ala Thr Leu Ile Val Asn Val Lys Pro Gln Ile Tyr Glu
            420                 425                 430
Lys Ala Val Ser Ser Phe Pro Asp Pro Ala Leu Tyr Pro Leu Gly Ser
        435                 440                 445
Arg Gln Ile Leu Thr Cys Thr Ala Tyr Gly Ile Pro Gln Pro Thr Ile
    450                 455                 460
Lys Trp Phe Trp His Pro Cys Asn His Asn His Ser Glu Ala Arg Cys
465                 470                 475                 480
Asp Phe Cys Ser Asn Asn Glu Glu Ser Phe Ile Leu Asp Ala Asp Ser
                485                 490                 495
Asn Met Gly Asn Arg Ile Glu Ser Ile Thr Gln Arg Met Ala Ile Ile
            500                 505                 510
Glu Gly Lys Asn Lys Met Ala Ser Thr Leu Val Val Ala Asp Ser Arg
        515                 520                 525
Ile Ser Gly Ile Tyr Ile Cys Ile Ala Ser Asn Lys Val Gly Thr Val
    530                 535                 540
Gly Arg Asn Ile Ser Phe Tyr Ile Thr Asp Val Pro Asn Gly Phe His
545                 550                 555                 560
```

```
Val Asn Leu Glu Lys Met Pro Thr Glu Gly Glu Asp Leu Lys Leu Ser
                565                 570                 575
Cys Thr Val Asn Lys Phe Leu Tyr Arg Asp Val Thr Trp Ile Leu Leu
            580                 585                 590
Arg Thr Val Asn Asn Arg Thr Met His Tyr Ser Ile Ser Lys Gln Lys
        595                 600                 605
Met Ala Ile Thr Lys Glu His Ser Ile Thr Leu Asn Leu Thr Ile Met
    610                 615                 620
Asn Val Ser Leu Gln Asp Ser Gly Thr Tyr Ala Cys Arg Ala Arg Asn
625                 630                 635                 640
Val Tyr Thr Gly Glu Glu Ile Leu Gln Lys Lys Glu Ile Thr Ile Arg
                645                 650                 655
Asp Gln Glu Ala Pro Tyr Leu Leu Arg Asn Leu Ser Asp His Thr Val
            660                 665                 670
Ala Ile Ser Ser Ser Thr Thr Leu Asp Cys His Ala Asn Gly Val Pro
        675                 680                 685
Glu Pro Gln Ile Thr Trp Phe Lys Asn Asn His Lys Ile Gln Gln Glu
    690                 695                 700
Pro Gly Ile Ile Leu Gly Pro Gly Ser Ser Thr Leu Phe Ile Glu Arg
705                 710                 715                 720
Val Thr Glu Glu Asp Glu Gly Val Tyr His Cys Lys Ala Thr Asn Gln
                725                 730                 735
Lys Gly Ser Val Glu Ser Ser Ala Tyr Leu Thr Val Gln Gly Thr Ser
            740                 745                 750
Asp Lys Ser Asn Leu Glu Leu Ile Thr Leu Thr Cys Thr Cys Val Ala
        755                 760                 765
Ala Thr Leu Phe Trp Leu Leu Leu Thr Leu Leu Ile Arg Lys Met Lys
    770                 775                 780
Arg Ser Ser Ser Glu Ile Lys Thr Asp Tyr Leu Ser Ile Ile Met Asp
785                 790                 795                 800
Pro Asp Glu Val Pro Leu Asp Glu Gln Cys Glu Arg Leu Pro Tyr Asp
                805                 810                 815
Ala Ser Lys Trp Glu Phe Ala Arg Glu Arg Leu Lys Leu Gly Lys Ser
            820                 825                 830
Leu Gly Arg Gly Ala Phe Gly Lys Val Val Gln Ala Ser Ala Phe Gly
        835                 840                 845
Ile Lys Lys Ser Pro Thr Cys Arg Thr Val Ala Val Lys Met Leu Lys
    850                 855                 860
Glu Gly Ala Thr Ala Ser Glu Tyr Lys Ala Leu Met Thr Glu Leu Lys
865                 870                 875                 880
Ile Leu Thr His Ile Gly His His Leu Asn Val Val Asn Leu Leu Gly
                885                 890                 895
Ala Cys Thr Lys Gln Gly Gly Pro Leu Met Val Ile Val Glu Tyr Cys
            900                 905                 910
Lys Tyr Gly Asn Leu Ser Asn Tyr Leu Lys Ser Lys Arg Asp Leu Phe
        915                 920                 925
Phe Leu Asn Lys Asp Ala Ala Leu His Met Glu Pro Lys Lys Glu Lys
    930                 935                 940
Met Glu Pro Gly Leu Glu Gln Gly Lys Lys Pro Arg Leu Asp Ser Val
945                 950                 955                 960
Thr Ser Ser Glu Ser Phe Ala Ser Ser Gly Phe Gln Glu Asp Lys Ser
                965                 970                 975
Leu Ser Asp Val Glu Glu Glu Glu Asp Ser Asp Gly Phe Tyr Lys Glu
            980                 985                 990
Pro Ile Thr Met Glu Asp Leu Ile  Ser Tyr Ser Phe Gln  Val Ala Arg
        995                 1000                1005
Gly Met  Glu Phe Leu Ser Ser  Arg Lys Cys Ile His  Arg Asp Leu
    1010                1015                1020
Ala Ala  Arg Asn Ile Leu Leu  Ser Glu Asn Asn Val  Val Lys Ile
    1025                1030                1035
Cys Asp  Phe Gly Leu Ala Arg  Asp Ile Tyr Lys Asn  Pro Asp Tyr
    1040                1045                1050
Val Arg  Lys Gly Asp Thr Arg  Leu Pro Leu Lys Trp  Met Ala Pro
```

```
    1055                1060                1065
Glu Ser  Ile Phe Asp Lys Ile  Tyr Ser Thr Lys Ser  Asp Val Trp
    1070                1075                1080
Ser Tyr  Gly Val Leu Leu Trp  Glu Ile Phe Ser Leu  Gly Gly Ser
    1085                1090                1095
Pro Tyr  Pro Gly Val Gln Met  Asp Glu Asp Phe Cys  Ser Arg Leu
    1100                1105                1110
Arg Glu  Gly Met Arg Met Arg  Ala Pro Glu Tyr Ser  Thr Pro Glu
    1115                1120                1125
Ile Tyr  Gln Ile Met Leu Asp  Cys Trp His Arg Asp  Pro Lys Glu
    1130                1135                1140
Arg Pro  Arg Phe Ala Glu Leu  Val Glu Lys Leu Gly  Asp Leu Leu
    1145                1150                1155
Gln Ala  Asn Val Gln Gln Asp  Gly Lys Asp Tyr Ile  Pro Ile Asn
    1160                1165                1170
Ala Ile  Leu Thr Gly Asn Ser  Gly Phe Thr Tyr Ser  Thr Pro Ala
    1175                1180                1185
Phe Ser  Glu Asp Phe Phe Lys  Glu Ser Ile Ser Ala  Pro Lys Phe
    1190                1195                1200
Asn Ser  Gly Ser Ser Asp Asp  Val Arg Tyr Val Asn  Ala Phe Lys
    1205                1210                1215
Phe Met  Ser Leu Glu Arg Ile  Lys Thr Phe Glu Glu  Leu Leu Pro
    1220                1225                1230
Asn Ala  Thr Ser Met Phe Asp  Asp Tyr Gln Gly Asp  Ser Ser Thr
    1235                1240                1245
Leu Leu  Ala Ser Pro Met Leu  Lys Arg Phe Thr Trp  Thr Asp Ser
    1250                1255                1260
Lys Pro  Lys Ala Ser Leu Lys  Ile Asp Leu Arg Val  Thr Ser Lys
    1265                1270                1275
Ser Lys  Glu Ser Gly Leu Ser  Asp Val Ser Arg Pro  Ser Phe Cys
    1280                1285                1290
His Ser  Ser Cys Gly His Val  Ser Glu Gly Lys Arg  Arg Phe Thr
    1295                1300                1305
Tyr Asp  His Ala Glu Leu Glu  Arg Lys Ile Ala Cys  Cys Ser Pro
    1310                1315                1320
Pro Pro  Asp Tyr Asn Ser Val  Val Leu Tyr Ser Thr  Pro Pro Ile
    1325                1330                1335
```

<210> 115
<211> 1356
<212> PRT
<213> Homo sapiens

<400> 115

```
Met Gln Ser Lys Val Leu Leu Ala Val Ala Leu Trp Leu Cys Val Glu
1               5                  10                  15
Thr Arg Ala Ala Ser Val Gly Leu Pro Ser Val Ser Leu Asp Leu Pro
            20                  25                  30
Arg Leu Ser Ile Gln Lys Asp Ile Leu Thr Ile Lys Ala Asn Thr Thr
        35                  40                  45
Leu Gln Ile Thr Cys Arg Gly Gln Arg Asp Leu Asp Trp Leu Trp Pro
    50                  55                  60
Asn Asn Gln Ser Gly Ser Glu Gln Arg Val Glu Val Thr Glu Cys Ser
65                  70                  75                  80
Asp Gly Leu Phe Cys Lys Thr Leu Thr Ile Pro Lys Val Ile Gly Asn
                85                  90                  95
Asp Thr Gly Ala Tyr Lys Cys Phe Tyr Arg Glu Thr Asp Leu Ala Ser
            100                 105                 110
Val Ile Tyr Val Tyr Val Gln Asp Tyr Arg Ser Pro Phe Ile Ala Ser
        115                 120                 125
Val Ser Asp Gln His Gly Val Val Tyr Ile Thr Glu Asn Lys Asn Lys
    130                 135                 140
```

```
Thr Val Val Ile Pro Cys Leu Gly Ser Ile Ser Asn Leu Asn Val Ser
145                 150                 155                 160
Leu Cys Ala Arg Tyr Pro Glu Lys Arg Phe Val Pro Asp Gly Asn Arg
                165                 170                 175
Ile Ser Trp Asp Ser Lys Lys Gly Phe Thr Ile Pro Ser Tyr Met Ile
            180                 185                 190
Ser Tyr Ala Gly Met Val Phe Cys Glu Ala Lys Ile Asn Asp Glu Ser
        195                 200                 205
Tyr Gln Ser Ile Met Tyr Ile Val Val Val Val Gly Tyr Arg Ile Tyr
    210                 215                 220
Asp Val Val Leu Ser Pro Ser His Gly Ile Glu Leu Ser Val Gly Glu
225                 230                 235                 240
Lys Leu Val Leu Asn Cys Thr Ala Arg Thr Glu Leu Asn Val Gly Ile
                245                 250                 255
Asp Phe Asn Trp Glu Tyr Pro Ser Ser Lys His Gln His Lys Lys Leu
            260                 265                 270
Val Asn Arg Asp Leu Lys Thr Gln Ser Gly Ser Glu Met Lys Lys Phe
        275                 280                 285
Leu Ser Thr Leu Thr Ile Asp Gly Val Thr Arg Ser Asp Gln Gly Leu
    290                 295                 300
Tyr Thr Cys Ala Ala Ser Ser Gly Leu Met Thr Lys Lys Asn Ser Thr
305                 310                 315                 320
Phe Val Arg Val His Glu Lys Pro Phe Val Ala Phe Gly Ser Gly Met
                325                 330                 335
Glu Ser Leu Val Glu Ala Thr Val Gly Glu Arg Val Arg Ile Pro Ala
            340                 345                 350
Lys Tyr Leu Gly Tyr Pro Pro Pro Glu Ile Lys Trp Tyr Lys Asn Gly
        355                 360                 365
Ile Pro Leu Glu Ser Asn His Thr Ile Lys Ala Gly His Val Leu Thr
    370                 375                 380
Ile Met Glu Val Ser Glu Arg Asp Thr Gly Asn Tyr Thr Val Ile Leu
385                 390                 395                 400
Thr Asn Pro Ile Ser Lys Glu Lys Gln Ser His Val Val Ser Leu Val
                405                 410                 415
Val Tyr Val Pro Pro Gln Ile Gly Glu Lys Ser Leu Ile Ser Pro Val
            420                 425                 430
Asp Ser Tyr Gln Tyr Gly Thr Thr Gln Thr Leu Thr Cys Thr Val Tyr
        435                 440                 445
Ala Ile Pro Pro Pro His His Ile His Trp Tyr Trp Gln Leu Glu Glu
    450                 455                 460
Glu Cys Ala Asn Glu Pro Ser Gln Ala Val Ser Val Thr Asn Pro Tyr
465                 470                 475                 480
Pro Cys Glu Glu Trp Arg Ser Val Glu Asp Phe Gln Gly Gly Asn Lys
                485                 490                 495
Ile Glu Val Asn Lys Asn Gln Phe Ala Leu Ile Glu Gly Lys Asn Lys
            500                 505                 510
Thr Val Ser Thr Leu Val Ile Gln Ala Ala Asn Val Ser Ala Leu Tyr
        515                 520                 525
Lys Cys Glu Ala Val Asn Lys Val Gly Arg Gly Glu Arg Val Ile Ser
    530                 535                 540
Phe His Val Thr Arg Gly Pro Glu Ile Thr Leu Gln Pro Asp Met Gln
545                 550                 555                 560
Pro Thr Glu Gln Glu Ser Val Ser Leu Trp Cys Thr Ala Asp Arg Ser
                565                 570                 575
Thr Phe Glu Asn Leu Thr Trp Tyr Lys Leu Gly Pro Gln Pro Leu Pro
            580                 585                 590
Ile His Val Gly Glu Leu Pro Thr Pro Val Cys Lys Asn Leu Asp Thr
        595                 600                 605
Leu Trp Lys Leu Asn Ala Thr Met Phe Ser Asn Ser Thr Asn Asp Ile
    610                 615                 620
Leu Ile Met Glu Leu Lys Asn Ala Ser Leu Gln Asp Gln Gly Asp Tyr
625                 630                 635                 640
Val Cys Leu Ala Gln Asp Arg Lys Thr Lys Lys Arg His Cys Val Val
```

```
                645                 650                 655
Arg Gln Leu Thr Val Leu Glu Arg Val Ala Pro Thr Ile Thr Gly Asn
            660                 665                 670
Leu Glu Asn Gln Thr Thr Ser Ile Gly Glu Ser Ile Glu Val Ser Cys
        675                 680                 685
Thr Ala Ser Gly Asn Pro Pro Pro Gln Ile Met Trp Phe Lys Asp Asn
    690                 695                 700
Glu Thr Leu Val Glu Asp Ser Gly Ile Val Leu Lys Asp Gly Asn Arg
705                 710                 715                 720
Asn Leu Thr Ile Arg Arg Val Arg Lys Glu Asp Glu Gly Leu Tyr Thr
                725                 730                 735
Cys Gln Ala Cys Ser Val Leu Gly Cys Ala Lys Val Glu Ala Phe Phe
            740                 745                 750
Ile Ile Glu Gly Ala Gln Glu Lys Thr Asn Leu Glu Ile Ile Ile Leu
        755                 760                 765
Val Gly Thr Ala Val Ile Ala Met Phe Phe Trp Leu Leu Leu Val Ile
    770                 775                 780
Ile Leu Arg Thr Val Lys Arg Ala Asn Gly Gly Glu Leu Lys Thr Gly
785                 790                 795                 800
Tyr Leu Ser Ile Val Met Asp Pro Asp Glu Leu Pro Leu Asp Glu His
                805                 810                 815
Cys Glu Arg Leu Pro Tyr Asp Ala Ser Lys Trp Glu Phe Pro Arg Asp
            820                 825                 830
Arg Leu Lys Leu Gly Lys Pro Leu Gly Arg Gly Ala Phe Gly Gln Val
        835                 840                 845
Ile Glu Ala Asp Ala Phe Gly Ile Asp Lys Thr Ala Thr Cys Arg Thr
    850                 855                 860
Val Ala Val Lys Met Leu Lys Glu Gly Ala Thr His Ser Glu His Arg
865                 870                 875                 880
Ala Leu Met Ser Glu Leu Lys Ile Leu Ile His Ile Gly His His Leu
                885                 890                 895
Asn Val Val Asn Leu Leu Gly Ala Cys Thr Lys Pro Gly Gly Pro Leu
            900                 905                 910
Met Val Ile Val Glu Phe Cys Lys Phe Gly Asn Leu Ser Thr Tyr Leu
        915                 920                 925
Arg Ser Lys Arg Asn Glu Phe Val Pro Tyr Lys Thr Lys Gly Ala Arg
    930                 935                 940
Phe Arg Gln Gly Lys Asp Tyr Val Gly Ala Ile Pro Val Asp Leu Lys
945                 950                 955                 960
Arg Arg Leu Asp Ser Ile Thr Ser Ser Gln Ser Ser Ala Ser Ser Gly
                965                 970                 975
Phe Val Glu Glu Lys Ser Leu Ser Asp Val Glu Glu Glu Glu Ala Pro
            980                 985                 990
Glu Asp Leu Tyr Lys Asp Phe Leu  Thr Leu Glu His Leu  Ile Cys Tyr
        995                 1000                1005
Ser Phe  Gln Val Ala Lys Gly  Met Glu Phe Leu Ala  Ser Arg Lys
    1010                1015                1020
Cys Ile  His Arg Asp Leu Ala  Ala Arg Asn Ile Leu  Leu Ser Glu
    1025                1030                1035
Lys Asn  Val Val Lys Ile Cys  Asp Phe Gly Leu Ala  Arg Asp Ile
    1040                1045                1050
Tyr Lys  Asp Pro Asp Tyr Val  Arg Lys Gly Asp Ala  Arg Leu Pro
    1055                1060                1065
Leu Lys  Trp Met Ala Pro Glu  Thr Ile Phe Asp Arg  Val Tyr Thr
    1070                1075                1080
Ile Gln  Ser Asp Val Trp Ser  Phe Gly Val Leu Leu  Trp Glu Ile
    1085                1090                1095
Phe Ser  Leu Gly Ala Ser Pro  Tyr Pro Gly Val Lys  Ile Asp Glu
    1100                1105                1110
Glu Phe  Cys Arg Arg Leu Lys  Glu Gly Thr Arg Met  Arg Ala Pro
    1115                1120                1125
Asp Tyr  Thr Thr Pro Glu Met  Tyr Gln Thr Met Leu  Asp Cys Trp
    1130                1135                1140
```

```
His Gly  Glu Pro Ser Gln Arg  Pro Thr Phe Ser Glu  Leu Val Glu
    1145                 1150                 1155
His Leu  Gly Asn Leu Leu Gln  Ala Asn Ala Gln Gln  Asp Gly Lys
    1160                 1165                 1170
Asp Tyr  Ile Val Leu Pro Ile  Ser Glu Thr Leu Ser  Met Glu Glu
    1175                 1180                 1185
Asp Ser  Gly Leu Ser Leu Pro  Thr Ser Pro Val Ser  Cys Met Glu
    1190                 1195                 1200
Glu Glu  Glu Val Cys Asp Pro  Lys Phe His Tyr Asp  Asn Thr Ala
    1205                 1210                 1215
Gly Ile  Ser Gln Tyr Leu Gln  Asn Ser Lys Arg Lys  Ser Arg Pro
    1220                 1225                 1230
Val Ser  Val Lys Thr Phe Glu  Asp Ile Pro Leu Glu  Glu Pro Glu
    1235                 1240                 1245
Val Lys  Val Ile Pro Asp Asp  Asn Gln Thr Asp Ser  Gly Met Val
    1250                 1255                 1260
Leu Ala  Ser Glu Glu Leu Lys  Thr Leu Glu Asp Arg  Thr Lys Leu
    1265                 1270                 1275
Ser Pro  Ser Phe Gly Gly Met  Val Pro Ser Lys Ser  Arg Glu Ser
    1280                 1285                 1290
Val Ala  Ser Glu Gly Ser Asn  Gln Thr Ser Gly Tyr  Gln Ser Gly
    1295                 1300                 1305
Tyr His  Ser Asp Asp Thr Asp  Thr Thr Val Tyr Ser  Ser Glu Glu
    1310                 1315                 1320
Ala Glu  Leu Leu Lys Leu Ile  Glu Ile Gly Val Gln  Thr Gly Ser
    1325                 1330                 1335
Thr Ala  Gln Ile Leu Gln Pro  Asp Ser Gly Thr Thr  Leu Ser Ser
    1340                 1345                 1350
Pro Pro  Val
    1355
```

<210> 116
<211> 1186
<212> PRT
<213> Homo sapiens

<400> 116

```
Leu Glu Glu Lys Lys Val Cys Gln Gly Thr Ser Asn Lys Leu Thr Gln
1               5                  10                  15
Leu Gly Thr Phe Glu Asp His Phe Leu Ser Leu Gln Arg Met Phe Asn
            20                  25                  30
Asn Cys Glu Val Val Leu Gly Asn Leu Glu Ile Thr Tyr Val Gln Arg
        35                  40                  45
Asn Tyr Asp Leu Ser Phe Leu Lys Thr Ile Gln Glu Val Ala Gly Tyr
    50                  55                  60
Val Leu Ile Ala Leu Asn Thr Val Glu Arg Ile Pro Leu Glu Asn Leu
65                  70                  75                  80
Gln Ile Ile Arg Gly Asn Met Tyr Tyr Glu Asn Ser Tyr Ala Leu Ala
                85                  90                  95
Val Leu Ser Asn Tyr Asp Ala Asn Lys Thr Gly Leu Lys Glu Leu Pro
            100                 105                 110
Met Arg Asn Leu Gln Glu Ile Leu His Gly Ala Val Arg Phe Ser Asn
        115                 120                 125
Asn Pro Ala Leu Cys Asn Val Glu Ser Ile Gln Trp Arg Asp Ile Val
    130                 135                 140
Ser Ser Asp Phe Leu Ser Asn Met Ser Met Asp Phe Gln Asn His Leu
145                 150                 155                 160
Gly Ser Cys Gln Lys Cys Asp Pro Ser Cys Pro Asn Gly Ser Cys Trp
                165                 170                 175
Gly Ala Gly Glu Glu Asn Cys Gln Lys Leu Thr Lys Ile Ile Cys Ala
            180                 185                 190
Gln Gln Cys Ser Gly Arg Cys Arg Gly Lys Ser Pro Ser Asp Cys Cys
```

```
        195                 200                 205
His Asn Gln Cys Ala Ala Gly Cys Thr Gly Pro Arg Glu Ser Asp Cys
    210                 215                 220
Leu Val Cys Arg Lys Phe Arg Asp Glu Ala Thr Cys Lys Asp Thr Cys
225                 230                 235                 240
Pro Pro Leu Met Leu Tyr Asn Pro Thr Thr Tyr Gln Met Asp Val Asn
                245                 250                 255
Pro Glu Gly Lys Tyr Ser Phe Gly Ala Thr Cys Val Lys Lys Cys Pro
            260                 265                 270
Arg Asn Tyr Val Val Thr Asp His Gly Ser Cys Val Arg Ala Cys Gly
        275                 280                 285
Ala Asp Ser Tyr Glu Met Glu Glu Asp Gly Val Arg Lys Cys Lys Lys
    290                 295                 300
Cys Glu Gly Pro Cys Arg Lys Val Cys Asn Gly Ile Gly Ile Gly Glu
305                 310                 315                 320
Phe Lys Asp Ser Leu Ser Ile Asn Ala Thr Asn Ile Lys His Phe Lys
                325                 330                 335
Asn Cys Thr Ser Ile Ser Gly Asp Leu His Ile Leu Pro Val Ala Phe
            340                 345                 350
Arg Gly Asp Ser Phe Thr His Thr Pro Pro Leu Asp Pro Gln Glu Leu
        355                 360                 365
Asp Ile Leu Lys Thr Val Lys Glu Ile Thr Gly Phe Leu Leu Ile Gln
    370                 375                 380
Ala Trp Pro Glu Asn Arg Thr Asp Leu His Ala Phe Glu Asn Leu Glu
385                 390                 395                 400
Ile Ile Arg Gly Arg Thr Lys Gln His Gly Gln Phe Ser Leu Ala Val
                405                 410                 415
Val Ser Leu Asn Ile Thr Ser Leu Gly Leu Arg Ser Leu Lys Glu Ile
            420                 425                 430
Ser Asp Gly Asp Val Ile Ile Ser Gly Asn Lys Asn Leu Cys Tyr Ala
        435                 440                 445
Asn Thr Ile Asn Trp Lys Lys Leu Phe Gly Thr Ser Gly Gln Lys Thr
    450                 455                 460
Lys Ile Ile Ser Asn Arg Gly Glu Asn Ser Cys Lys Ala Thr Gly Gln
465                 470                 475                 480
Val Cys His Ala Leu Cys Ser Pro Glu Gly Cys Trp Gly Pro Glu Pro
                485                 490                 495
Arg Asp Cys Val Ser Cys Arg Asn Val Ser Arg Gly Arg Glu Cys Val
            500                 505                 510
Asp Lys Cys Asn Leu Leu Glu Gly Glu Pro Arg Glu Phe Val Glu Asn
        515                 520                 525
Ser Glu Cys Ile Gln Cys His Pro Glu Cys Leu Pro Gln Ala Met Asn
    530                 535                 540
Ile Thr Cys Thr Gly Arg Gly Pro Asp Asn Cys Ile Gln Cys Ala His
545                 550                 555                 560
Tyr Ile Asp Gly Pro His Cys Val Lys Thr Cys Pro Ala Gly Val Met
                565                 570                 575
Gly Glu Asn Asn Thr Leu Val Trp Lys Tyr Ala Asp Ala Gly His Val
            580                 585                 590
Cys His Leu Cys His Pro Asn Cys Thr Tyr Gly Cys Thr Gly Pro Gly
        595                 600                 605
Leu Glu Gly Cys Pro Thr Asn Gly Pro Lys Ile Pro Ser Ile Ala Thr
    610                 615                 620
Gly Met Val Gly Ala Leu Leu Leu Leu Leu Val Val Ala Leu Gly Ile
625                 630                 635                 640
Gly Leu Phe Met Arg Arg Arg His Ile Val Arg Lys Arg Thr Leu Arg
                645                 650                 655
Arg Leu Leu Gln Glu Arg Glu Leu Val Glu Pro Leu Thr Pro Ser Gly
            660                 665                 670
Glu Ala Pro Asn Gln Ala Leu Leu Arg Ile Leu Lys Glu Thr Glu Phe
        675                 680                 685
Lys Lys Ile Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys
    690                 695                 700
```

```
Gly Leu Trp Ile Pro Glu Gly Glu Lys Val Lys Ile Pro Val Ala Ile
705                 710                 715                 720
Lys Glu Leu Arg Glu Ala Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu
                725                 730                 735
Asp Glu Ala Tyr Val Met Ala Ser Val Asp Asn Pro His Val Cys Arg
            740                 745                 750
Leu Leu Gly Ile Cys Leu Thr Ser Thr Val Gln Leu Ile Thr Gln Leu
        755                 760                 765
Met Pro Phe Gly Cys Leu Leu Asp Tyr Val Arg Glu His Lys Asp Asn
    770                 775                 780
Ile Gly Ser Gln Tyr Leu Leu Asn Trp Cys Val Gln Ile Ala Lys Gly
785                 790                 795                 800
Met Asn Tyr Leu Glu Asp Arg Arg Leu Val His Arg Asp Leu Ala Ala
                805                 810                 815
Arg Asn Val Leu Val Lys Thr Pro Gln His Val Lys Ile Thr Asp Phe
            820                 825                 830
Gly Leu Ala Lys Leu Leu Gly Ala Glu Glu Lys Glu Tyr His Ala Glu
        835                 840                 845
Gly Gly Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu His
    850                 855                 860
Arg Ile Tyr Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val
865                 870                 875                 880
Trp Glu Leu Met Thr Phe Gly Ser Lys Pro Tyr Asp Gly Ile Pro Ala
                885                 890                 895
Ser Glu Ile Ser Ser Ile Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro
            900                 905                 910
Pro Ile Cys Thr Ile Asp Val Tyr Met Ile Met Val Lys Cys Trp Met
        915                 920                 925
Ile Asp Ala Asp Ser Arg Pro Lys Phe Arg Glu Leu Ile Ile Glu Phe
    930                 935                 940
Ser Lys Met Ala Arg Asp Pro Gln Arg Tyr Leu Val Ile Gln Gly Asp
945                 950                 955                 960
Glu Arg Met His Leu Pro Ser Pro Thr Asp Ser Asn Phe Tyr Arg Ala
                965                 970                 975
Leu Met Asp Glu Glu Asp Met Asp Asp Val Val Asp Ala Asp Glu Tyr
            980                 985                 990
Leu Ile Pro Gln Gln Gly Phe Phe  Ser Ser Pro Ser Thr  Ser Arg Thr
        995                 1000                1005
Pro Leu  Leu Ser Ser Leu Ser  Ala Thr Ser Asn Asn  Ser Thr Val
    1010                1015                1020
Ala Cys  Ile Asp Arg Asn Gly  Leu Gln Ser Cys Pro  Ile Lys Glu
    1025                1030                1035
Asp Ser  Phe Leu Gln Arg Tyr  Ser Ser Asp Pro Thr  Gly Ala Leu
    1040                1045                1050
Thr Glu  Asp Ser Ile Asp Asp  Thr Phe Leu Pro Val  Pro Glu Tyr
    1055                1060                1065
Ile Asn  Gln Ser Val Pro Lys  Arg Pro Ala Gly Ser  Val Gln Asn
    1070                1075                1080
Pro Val  Tyr His Asn Gln Pro  Leu Asn Pro Ala Pro  Ser Arg Asp
    1085                1090                1095
Pro His  Tyr Gln Asp Pro His  Ser Thr Ala Val Gly  Asn Pro Glu
    1100                1105                1110
Tyr Leu  Asn Thr Val Gln Pro  Thr Cys Val Asn Ser  Thr Phe Asp
    1115                1120                1125
Ser Pro  Ala His Trp Ala Gln  Lys Gly Ser His Gln  Ile Ser Leu
    1130                1135                1140
Asp Asn  Pro Asp Tyr Gln Gln  Asp Phe Phe Pro Lys  Glu Ala Lys
    1145                1150                1155
Pro Asn  Gly Ile Phe Lys Gly  Ser Thr Ala Glu Asn  Ala Glu Tyr
    1160                1165                1170
Leu Arg  Val Ala Pro Gln Ser  Ser Glu Phe Ile Gly  Ala
    1175                1180                1185
```

<210> 117
<211> 422
<212> PRT
<213> Homo sapiens

<400> 117
Met Asp Val Leu Ser Pro Gly Gln Gly Asn Asn Thr Thr Ser Pro Pro
1 5 10 15
Ala Pro Phe Glu Thr Gly Gly Asn Thr Thr Gly Ile Ser Asp Val Thr
20 25 30
Val Ser Tyr Gln Val Ile Thr Ser Leu Leu Leu Gly Thr Leu Ile Phe
35 40 45
Cys Ala Val Leu Gly Asn Ala Cys Val Val Ala Ala Ile Ala Leu Glu
50 55 60
Arg Ser Leu Gln Asn Val Ala Asn Tyr Leu Ile Gly Ser Leu Ala Val
65 70 75 80
Thr Asp Leu Met Val Ser Val Leu Val Leu Pro Met Ala Ala Leu Tyr
85 90 95
Gln Val Leu Asn Lys Trp Thr Leu Gly Gln Val Thr Cys Asp Leu Phe
100 105 110
Ile Ala Leu Asp Val Leu Cys Cys Thr Ser Ser Ile Leu His Leu Cys
115 120 125
Ala Ile Ala Leu Asp Arg Tyr Trp Ala Ile Thr Asp Pro Ile Asp Tyr
130 135 140
Val Asn Lys Arg Thr Pro Arg Arg Ala Ala Ala Leu Ile Ser Leu Thr
145 150 155 160
Trp Leu Ile Gly Phe Leu Ile Ser Ile Pro Pro Met Leu Gly Trp Arg
165 170 175
Thr Pro Glu Asp Arg Ser Asp Pro Asp Ala Cys Thr Ile Ser Lys Asp
180 185 190
His Gly Tyr Thr Ile Tyr Ser Thr Phe Gly Ala Phe Tyr Ile Pro Leu
195 200 205
Leu Leu Met Leu Val Leu Tyr Gly Arg Ile Phe Arg Ala Ala Arg Phe
210 215 220
Arg Ile Arg Lys Thr Val Lys Lys Val Glu Lys Thr Gly Ala Asp Thr
225 230 235 240
Arg His Gly Ala Ser Pro Ala Pro Gln Pro Lys Lys Ser Val Asn Gly
245 250 255
Glu Ser Gly Ser Arg Asn Trp Arg Leu Gly Val Glu Ser Lys Ala Gly
260 265 270
Gly Ala Leu Cys Ala Asn Gly Ala Val Arg Gln Gly Asp Asp Gly Ala
275 280 285
Ala Leu Glu Val Ile Glu Val His Arg Val Gly Asn Ser Lys Glu His
290 295 300
Leu Pro Leu Pro Ser Glu Ala Gly Pro Thr Pro Cys Ala Pro Ala Ser
305 310 315 320
Phe Glu Arg Lys Asn Glu Arg Asn Ala Glu Ala Lys Arg Lys Met Ala
325 330 335
Leu Ala Arg Glu Arg Lys Thr Val Lys Thr Leu Gly Ile Ile Met Gly
340 345 350
Thr Phe Ile Leu Cys Trp Leu Pro Phe Phe Ile Val Ala Leu Val Leu
355 360 365
Pro Phe Cys Glu Ser Ser Cys His Met Pro Thr Leu Leu Gly Ala Ile
370 375 380
Ile Asn Trp Leu Gly Tyr Ser Asn Ser Leu Leu Asn Pro Val Ile Tyr
385 390 395 400
Ala Tyr Phe Asn Lys Asp Phe Gln Asn Ala Phe Lys Lys Ile Ile Lys
405 410 415
Cys Lys Phe Cys Arg Gln
420

<210> 118
<211> 129
<212> PRT
<213> Homo sapiens

<400> 118
His Lys Cys Asp Ile Thr Leu Gln Glu Ile Ile Lys Thr Leu Asn Ser
1 5 10 15
Leu Thr Glu Gln Lys Thr Leu Cys Thr Glu Leu Thr Val Thr Asp Ile
20 25 30
Phe Ala Ala Ser Lys Asn Thr Thr Glu Lys Glu Thr Phe Cys Arg Ala
35 40 45
Ala Thr Val Leu Arg Gln Phe Tyr Ser His His Glu Lys Asp Thr Arg
50 55 60
Cys Leu Gly Ala Thr Ala Gln Gln Phe His Arg His Lys Gln Leu Ile
65 70 75 80
Arg Phe Leu Lys Arg Leu Asp Arg Asn Leu Trp Gly Leu Ala Gly Leu
85 90 95
Asn Ser Cys Pro Val Lys Glu Ala Asn Gln Ser Thr Leu Glu Asn Phe
100 105 110
Leu Glu Arg Leu Lys Thr Ile Met Arg Glu Lys Tyr Ser Lys Cys Ser
115 120 125
Ser

<210> 119
<211> 113
<212> PRT
<213> Homo sapiens

<400> 119
Met Gly Pro Val Pro Pro Ser Thr Ala Leu Arg Glu Leu Ile Glu Glu
1 5 10 15
Leu Val Asn Ile Thr Gln Asn Gln Lys Ala Pro Leu Cys Asn Gly Ser
20 25 30
Met Val Trp Ser Ile Asn Leu Thr Ala Gly Met Tyr Cys Ala Ala Leu
35 40 45
Glu Ser Leu Ile Asn Val Ser Gly Cys Ser Ala Ile Glu Lys Thr Gln
50 55 60
Arg Met Leu Ser Gly Phe Cys Pro His Lys Val Ser Ala Gly Gln Phe
65 70 75 80
Ser Ser Leu His Val Arg Asp Thr Lys Ile Glu Val Ala Gln Phe Val
85 90 95
Lys Asp Leu Leu Leu His Leu Lys Lys Leu Phe Arg Glu Gly Arg Phe
100 105 110
Asn

<210> 120
<211> 726
<212> PRT
<213> Homo sapiens

<400> 120
Val Thr Lys Leu Leu Pro Ala Leu Leu Leu Gln His Val Leu Leu His
1 5 10 15
Leu Leu Leu Leu Pro Ile Ala Ile Pro Tyr Ala Glu Gly Gln Arg Lys
20 25 30
Arg Arg Asn Thr Ile His Glu Phe Lys Lys Ser Ala Lys Thr Thr Leu
35 40 45
Ile Lys Ile Asp Pro Ala Leu Lys Ile Lys Thr Lys Lys Val Asn Thr
50 55 60
Ala Asp Gln Cys Ala Asn Arg Cys Thr Arg Asn Lys Gly Leu Pro Phe

65 70 75 80
Thr Cys Lys Ala Phe Val Phe Asp Lys Ala Arg Lys Gln Cys Leu Trp
85 90 95
Phe Pro Phe Asn Ser Met Ser Ser Gly Val Lys Lys Glu Phe Gly His
100 105 110
Glu Phe Asp Leu Tyr Glu Asn Lys Asp Tyr Ile Arg Asn Cys Ile Ile
115 120 125
Gly Lys Gly Arg Ser Tyr Lys Gly Thr Val Ser Ile Thr Lys Ser Gly
130 135 140
Ile Lys Cys Gln Pro Trp Ser Ser Met Ile Pro His Glu His Ser Phe
145 150 155 160
Leu Pro Ser Ser Tyr Arg Gly Lys Asp Leu Gln Glu Asn Tyr Cys Arg
165 170 175
Asn Pro Arg Gly Glu Glu Gly Gly Pro Trp Cys Phe Thr Ser Asn Pro
180 185 190
Glu Val Arg Tyr Glu Val Cys Asp Ile Pro Gln Cys Ser Glu Val Glu
195 200 205
Cys Met Thr Cys Asn Gly Glu Ser Tyr Arg Gly Leu Met Asp His Thr
210 215 220
Glu Ser Gly Lys Ile Cys Gln Arg Trp Asp His Gln Thr Pro His Arg
225 230 235 240
His Lys Phe Leu Pro Glu Arg Tyr Pro Asp Lys Gly Phe Asp Asp Asn
245 250 255
Tyr Cys Arg Asn Pro Asp Gly Gln Pro Arg Pro Trp Cys Tyr Thr Leu
260 265 270
Asp Pro His Thr Arg Trp Glu Tyr Cys Ala Ile Lys Thr Cys Ala Asp
275 280 285
Asn Thr Met Asn Asp Thr Asp Val Pro Leu Glu Thr Thr Glu Cys Ile
290 295 300
Gln Gly Gln Gly Glu Gly Tyr Arg Gly Thr Val Asn Thr Ile Trp Asn
305 310 315 320
Gly Ile Pro Cys Gln Arg Trp Asp Ser Gln Tyr Pro His Glu His Asp
325 330 335
Met Thr Pro Glu Asn Phe Lys Cys Lys Asp Leu Arg Glu Asn Tyr Cys
340 345 350
Arg Asn Pro Asp Gly Ser Glu Ser Pro Trp Cys Phe Thr Thr Asp Pro
355 360 365
Asn Ile Arg Val Gly Tyr Cys Ser Gln Ile Pro Asn Cys Asp Met Ser
370 375 380
His Gly Gln Asp Cys Tyr Arg Gly Asn Gly Lys Asn Tyr Met Gly Asn
385 390 395 400
Leu Ser Gln Thr Arg Ser Gly Leu Thr Cys Ser Met Trp Asp Lys Asn
405 410 415
Met Glu Asp Leu His Arg His Ile Phe Trp Glu Pro Asp Ala Ser Lys
420 425 430
Leu Asn Glu Asn Tyr Cys Arg Asn Pro Asp Asp Asp Ala His Gly Pro
435 440 445
Trp Cys Tyr Thr Gly Asn Pro Leu Ile Pro Trp Asp Tyr Cys Pro Ile
450 455 460
Ser Arg Cys Glu Gly Asp Thr Thr Pro Thr Ile Val Asn Leu Asp His
465 470 475 480
Pro Val Ile Ser Cys Ala Lys Thr Lys Gln Leu Arg Val Val Asn Gly
485 490 495
Ile Pro Thr Arg Thr Asn Ile Gly Trp Met Val Ser Leu Arg Tyr Arg
500 505 510
Asn Lys His Ile Cys Gly Gly Ser Leu Ile Lys Glu Ser Trp Val Leu
515 520 525
Thr Ala Arg Gln Cys Phe Pro Ser Arg Asp Leu Lys Asp Tyr Glu Ala
530 535 540
Trp Leu Gly Ile His Asp Val His Gly Arg Gly Asp Glu Lys Cys Lys
545 550 555 560
Gln Val Leu Asn Val Ser Gln Leu Val Tyr Gly Pro Glu Gly Ser Asp
565 570 575

Leu Val Leu Met Lys Leu Ala Arg Pro Ala Val Leu Asp Asp Phe Val
580 585 590
Ser Thr Ile Asp Leu Pro Asn Tyr Gly Cys Thr Ile Pro Glu Lys Thr
595 600 605
Ser Cys Ser Val Tyr Gly Trp Gly Tyr Thr Gly Leu Ile Asn Tyr Asp
610 615 620
Gly Leu Leu Arg Val Ala His Leu Tyr Ile Met Gly Asn Glu Lys Cys
625 630 635 640
Ser Gln His His Arg Gly Lys Val Thr Leu Asn Glu Ser Glu Ile Cys
645 650 655
Ala Gly Ala Glu Lys Ile Gly Ser Gly Pro Cys Glu Gly Asp Tyr Gly
660 665 670
Gly Pro Leu Val Cys Glu Gln His Lys Met Arg Met Val Leu Gly Val
675 680 685
Ile Val Pro Gly Arg Gly Cys Ala Ile Pro Asn Arg Pro Gly Ile Phe
690 695 700
Val Arg Val Ala Tyr Tyr Ala Lys Trp Ile His Lys Ile Ile Leu Thr
705 710 715 720
Tyr Lys Val Pro Gln Ser
725

<210> 121
<211> 191
<212> PRT
<213> Homo sapiens

<400> 121
Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg
1 5 10 15
Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu
20 25 30
Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro
35 40 45
Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg
50 55 60
Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu
65 70 75 80
Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val
85 90 95
Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp
100 105 110

Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu
115 120 125
Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser
130 135 140
Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr
145 150 155 160
Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe
165 170 175
Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe
180 185 190

<210> 122
<211> 156
<212> PRT
<213> Homo sapiens

<400> 122
Ala Tyr Arg Pro Ser Glu Thr Leu Cys Gly Gly Glu Leu Val Asp Thr
1 5 10 15
Leu Gln Phe Val Cys Gly Asp Arg Gly Phe Tyr Phe Ser Arg Pro Ala

```
            20                  25                  30
Ser Arg Val Ser Arg Arg Ser Arg Gly Ile Val Glu Glu Cys Cys Phe
        35                  40                  45
Arg Ser Cys Asp Leu Ala Leu Leu Glu Thr Tyr Cys Ala Thr Pro Ala
    50                  55                  60
Lys Ser Glu Arg Asp Val Ser Thr Pro Pro Thr Val Leu Pro Asp Asn
65                  70                  75                  80
Phe Pro Arg Tyr Pro Val Gly Lys Phe Phe Gln Tyr Asp Thr Trp Lys
                85                  90                  95
Gln Ser Thr Gln Arg Leu Arg Arg Gly Leu Pro Ala Leu Leu Arg Ala
            100                 105                 110
Arg Arg Gly His Val Leu Ala Lys Glu Leu Glu Ala Phe Arg Glu Ala
        115                 120                 125
Lys Arg His Arg Pro Leu Ile Ala Leu Pro Thr Gln Asp Pro Ala His
    130                 135                 140
Gly Gly Ala Pro Pro Glu Met Ala Ser Asn Arg Lys
145                 150                 155


<210>  123
<211>  735
<212>  PRT
<213>  Homo sapiens

<400>  123
Glu Val Lys Gln Glu Asn Arg Leu Leu Asn Glu Ser Glu Ser Ser Ser
1               5                   10                  15
Gln Gly Leu Leu Gly Tyr Tyr Phe Ser Asp Leu Asn Phe Gln Ala Pro
            20                  25                  30
Met Val Val Thr Ser Ser Thr Thr Gly Asp Leu Ser Ile Pro Ser Ser
        35                  40                  45
Glu Leu Glu Asn Ile Pro Ser Glu Asn Gln Tyr Phe Gln Ser Ala Ile
    50                  55                  60
Trp Ser Gly Phe Ile Lys Val Lys Lys Ser Asp Glu Tyr Thr Phe Ala
65                  70                  75                  80
Thr Ser Ala Asp Asn His Val Thr Met Trp Val Asp Asp Gln Glu Val
                85                  90                  95
Ile Asn Lys Ala Ser Asn Ser Asn Lys Ile Arg Leu Glu Lys Gly Arg
            100                 105                 110
Leu Tyr Gln Ile Lys Ile Gln Tyr Gln Arg Glu Asn Pro Thr Glu Lys
        115                 120                 125
Gly Leu Asp Phe Lys Leu Tyr Trp Thr Asp Ser Gln Asn Lys Lys Glu
    130                 135                 140
Val Ile Ser Ser Asp Asn Leu Gln Leu Pro Glu Leu Lys Gln Lys Ser
145                 150                 155                 160
Ser Asn Ser Arg Lys Lys Arg Ser Thr Ser Ala Gly Pro Thr Val Pro
                165                 170                 175
Asp Arg Asp Asn Asp Gly Ile Pro Asp Ser Leu Glu Val Glu Gly Tyr
            180                 185                 190
Thr Val Asp Val Lys Asn Lys Arg Thr Phe Leu Ser Pro Trp Ile Ser
        195                 200                 205
Asn Ile His Glu Lys Lys Gly Leu Thr Lys Tyr Lys Ser Ser Pro Glu
    210                 215                 220
Lys Trp Ser Thr Ala Ser Asp Pro Tyr Ser Asp Phe Glu Lys Val Thr
225                 230                 235                 240
Gly Arg Ile Asp Lys Asn Val Ser Pro Glu Ala Arg His Pro Leu Val
                245                 250                 255
Ala Ala Tyr Pro Ile Val His Val Asp Met Glu Asn Ile Ile Leu Ser
            260                 265                 270
Lys Asn Glu Asp Gln Ser Thr Gln Asn Thr Asp Ser Gln Thr Arg Thr
        275                 280                 285
Ile Ser Lys Asn Thr Ser Thr Ser Arg Thr His Thr Ser Glu Val His
    290                 295                 300
```

Gly Asn Ala Glu Val His Ala Ser Phe Phe Asp Ile Gly Gly Ser Val
305 310 315 320
Ser Ala Gly Phe Ser Asn Ser Asn Ser Ser Thr Val Ala Ile Asp His
325 330 335
Ser Leu Ser Leu Ala Gly Glu Arg Thr Trp Ala Glu Thr Met Gly Leu
340 345 350
Asn Thr Ala Asp Thr Ala Arg Leu Asn Ala Asn Ile Arg Tyr Val Asn
355 360 365
Thr Gly Thr Ala Pro Ile Tyr Asn Val Leu Pro Thr Thr Ser Leu Val
370 375 380
Leu Gly Lys Asn Gln Thr Leu Ala Thr Ile Lys Ala Lys Glu Asn Gln
385 390 395 400
Leu Ser Gln Ile Leu Ala Pro Asn Asn Tyr Tyr Pro Ser Lys Asn Leu
405 410 415
Ala Pro Ile Ala Leu Asn Ala Gln Asp Asp Phe Ser Ser Thr Pro Ile
420 425 430
Thr Met Asn Tyr Asn Gln Phe Leu Glu Leu Glu Lys Thr Lys Gln Leu
435 440 445
Arg Leu Asp Thr Asp Gln Val Tyr Gly Asn Ile Ala Thr Tyr Asn Phe
450 455 460
Glu Asn Gly Arg Val Arg Val Asp Thr Gly Ser Asn Trp Ser Glu Val
465 470 475 480
Leu Pro Gln Ile Gln Glu Thr Thr Ala Arg Ile Ile Phe Asn Gly Lys
485 490 495
Asp Leu Asn Leu Val Glu Arg Arg Ile Ala Ala Val Asn Pro Ser Asp
500 505 510
Pro Leu Glu Thr Thr Lys Pro Asp Met Thr Leu Lys Glu Ala Leu Lys
515 520 525
Ile Ala Phe Gly Phe Asn Glu Pro Asn Gly Asn Leu Gln Tyr Gln Gly
530 535 540
Lys Asp Ile Thr Glu Phe Asp Phe Asn Phe Asp Gln Gln Thr Ser Gln
545 550 555 560
Asn Ile Lys Asn Gln Leu Ala Glu Leu Asn Ala Thr Asn Ile Tyr Thr
565 570 575
Val Leu Asp Lys Ile Lys Leu Asn Ala Lys Met Asn Ile Leu Ile Arg
580 585 590
Asp Lys Arg Phe His Tyr Asp Arg Asn Asn Ile Ala Val Gly Ala Asp
595 600 605
Glu Ser Val Val Lys Glu Ala His Arg Glu Val Ile Asn Ser Ser Thr
610 615 620
Glu Gly Leu Leu Leu Asn Ile Asp Lys Asp Ile Arg Lys Ile Leu Ser
625 630 635 640
Gly Tyr Ile Val Glu Ile Glu Asp Thr Glu Gly Leu Lys Glu Val Ile
645 650 655
Asn Asp Arg Tyr Asp Met Leu Asn Ile Ser Ser Leu Arg Gln Asp Gly
660 665 670
Lys Thr Phe Ile Asp Phe Lys Lys Tyr Asn Asp Lys Leu Pro Leu Tyr
675 680 685
Ile Ser Asn Pro Asn Tyr Lys Val Asn Val Tyr Ala Val Thr Lys Glu
690 695 700
Asn Thr Ile Ile Asn Pro Ser Glu Asn Gly Asp Thr Ser Thr Asn Gly
705 710 715 720
Ile Lys Lys Ile Leu Ile Phe Ser Lys Lys Gly Tyr Glu Ile Gly
725 730 735

<210> 124
<211> 509
<212> PRT
<213> Homo sapiens

<400> 124
Met Lys Val Lys Gly Thr Arg Arg Asn Tyr Gln His Leu Trp Arg Trp

```
1               5                   10                  15
Gly Thr Leu Leu Leu Gly Met Leu Met Ile Cys Ser Ala Thr Glu Lys
            20                  25                  30
Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
        35                  40                  45
Thr Thr Leu Phe Cys Ala Ser Asp Ala Arg Ala Tyr Asp Thr Glu Val
    50                  55                  60
His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65                  70                  75                  80
Gln Glu Val Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys
                85                  90                  95
Asn Asn Met Val Glu Gln Met Gln Glu Asp Ile Ile Ser Leu Trp Asp
            100                 105                 110
Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
        115                 120                 125
Asn Cys Thr Asp Leu Gly Lys Ala Thr Asn Thr Asn Ser Ser Asn Trp
    130                 135                 140
Lys Glu Glu Ile Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Thr
145                 150                 155                 160
Thr Ser Ile Arg Asp Lys Ile Gln Lys Glu Asn Ala Leu Phe Arg Asn
                165                 170                 175
Leu Asp Val Val Pro Ile Asp Asn Ala Ser Thr Thr Thr Asn Tyr Thr
            180                 185                 190
Asn Tyr Arg Leu Ile His Cys Asn Arg Ser Val Ile Thr Gln Ala Cys
        195                 200                 205
Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Thr Pro Ala
    210                 215                 220
Gly Phe Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Asn Gly Lys Gly
225                 230                 235                 240
Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro
                245                 250                 255
Ile Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu
            260                 265                 270
Val Val Ile Arg Ser Asp Asn Phe Thr Asn Asn Ala Lys Thr Ile Ile
        275                 280                 285
Val Gln Leu Asn Glu Ser Val Ala Ile Asn Cys Thr Arg Pro Asn Asn
    290                 295                 300
Asn Thr Arg Lys Ser Ile Tyr Ile Gly Pro Gly Arg Ala Phe His Thr
305                 310                 315                 320
Thr Gly Arg Ile Ile Gly Asp Ile Arg Lys Ala His Cys Asn Ile Ser
                325                 330                 335
Arg Ala Gln Trp Asn Asn Thr Leu Glu Gln Ile Val Lys Lys Leu Arg
            340                 345                 350
Glu Gln Phe Gly Asn Asn Lys Thr Ile Val Phe Asn Gln Ser Ser Gly
        355                 360                 365
Gly Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Arg Gly Glu Phe
    370                 375                 380
Phe Tyr Cys Asn Thr Thr Gln Leu Phe Asn Asn Thr Trp Arg Leu Asn
385                 390                 395                 400
His Thr Glu Gly Thr Lys Gly Asn Asp Thr Ile Ile Leu Pro Cys Arg
                405                 410                 415
Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Lys Ala Met Tyr
            420                 425                 430
Ala Pro Pro Ile Gly Gly Gln Ile Ser Cys Ser Ser Asn Ile Thr Gly
        435                 440                 445
Leu Leu Leu Thr Arg Asp Gly Gly Thr Asn Val Thr Asn Asp Thr Glu
    450                 455                 460
Val Phe Arg Pro Gly Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu
465                 470                 475                 480
Leu Tyr Lys Tyr Lys Val Ile Lys Ile Glu Pro Leu Gly Ile Ala Pro
                485                 490                 495
Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg
            500                 505
```

<210> 125
<211> 101
<212> PRT
<213> Homo sapiens

<400> 125
Ser Trp Val Ile Pro Pro Ile Ser Cys Pro Glu Asn Glu Lys Gly Pro
1 5 10 15
Phe Pro Lys Asn Leu Val Gln Ile Lys Ser Asn Lys Asp Lys Glu Gly
20 25 30
Lys Val Phe Tyr Ser Ile Thr Gly Gln Gly Ala Asp Thr Pro Pro Val
35 40 45
Gly Val Phe Ile Ile Glu Arg Glu Thr Gly Trp Leu Lys Val Thr Glu
50 55 60
Pro Leu Asp Arg Glu Arg Ile Ala Thr Tyr Thr Leu Phe Ser His Ala
65 70 75 80
Val Ser Ser Asn Gly Asn Ala Val Glu Asp Pro Met Glu Ile Leu Ile
85 90 95
Thr Val Thr Asp Gln
100

<210> 126
<211> 459
<212> PRT
<213> Homo sapiens

<400> 126
Glu Ile Cys Gly Pro Gly Ile Asp Ile Arg Asn Asp Tyr Gln Gln Leu
1 5 10 15
Lys Arg Leu Glu Asn Cys Thr Val Ile Glu Gly Tyr Leu His Ile Leu
20 25 30
Leu Ile Ser Lys Ala Glu Asp Tyr Arg Ser Tyr Arg Phe Pro Lys Leu
35 40 45
Thr Val Ile Thr Glu Tyr Leu Leu Leu Phe Arg Val Ala Gly Leu Glu
50 55 60
Ser Leu Gly Asp Leu Phe Pro Asn Leu Thr Val Ile Arg Gly Trp Lys
65 70 75 80
Leu Phe Tyr Asn Tyr Ala Leu Val Ile Phe Glu Met Thr Asn Leu Lys
85 90 95
Asp Ile Gly Leu Tyr Asn Leu Arg Asn Ile Thr Arg Gly Ala Ile Arg
100 105 110
Ile Glu Lys Asn Ala Asp Leu Cys Tyr Leu Ser Thr Val Asp Trp Ser
115 120 125
Leu Ile Leu Asp Ala Val Ser Asn Asn Tyr Ile Val Gly Asn Lys Pro
130 135 140
Pro Lys Glu Cys Gly Asp Leu Cys Pro Gly Thr Met Glu Glu Lys Pro
145 150 155 160
Met Cys Glu Lys Thr Thr Ile Asn Asn Glu Tyr Asn Tyr Arg Cys Trp
165 170 175
Thr Thr Asn Arg Cys Gln Lys Met Cys Pro Ser Thr Cys Gly Lys Arg
180 185 190
Ala Cys Thr Glu Asn Asn Glu Cys Cys His Pro Glu Cys Leu Gly Ser
195 200 205
Cys Ser Ala Pro Asp Asn Asp Thr Ala Cys Val Ala Cys Arg His Tyr
210 215 220
Tyr Tyr Ala Gly Val Cys Val Pro Ala Cys Pro Pro Asn Thr Tyr Arg
225 230 235 240
Phe Glu Gly Trp Arg Cys Val Asp Arg Asp Phe Cys Ala Asn Ile Leu
245 250 255
Ser Ala Glu Ser Ser Asp Ser Glu Gly Phe Val Ile His Asp Gly Glu

```
            260                 265                 270
Cys Met Gln Glu Cys Pro Ser Gly Phe Ile Arg Asn Gly Ser Gln Ser
        275                 280                 285
Met Tyr Cys Ile Pro Cys Glu Gly Pro Cys Pro Lys Val Cys Glu Glu
    290                 295                 300
Glu Lys Lys Thr Lys Thr Ile Asp Ser Val Thr Ser Ala Gln Met Leu
305                 310                 315                 320
Gln Gly Cys Thr Ile Phe Lys Gly Asn Leu Leu Ile Asn Ile Arg Arg
                325                 330                 335
Gly Asn Asn Ile Ala Ser Glu Leu Glu Asn Phe Met Gly Leu Ile Glu
            340                 345                 350
Val Val Thr Gly Tyr Val Lys Ile Arg His Ser His Ala Leu Val Ser
        355                 360                 365
Leu Ser Phe Leu Lys Asn Leu Arg Leu Ile Leu Gly Glu Glu Gln Leu
    370                 375                 380
Glu Gly Asn Tyr Ser Phe Tyr Val Leu Asp Asn Gln Asn Leu Gln Gln
385                 390                 395                 400
Leu Trp Asp Trp Asp His Arg Asn Leu Thr Ile Lys Ala Gly Lys Met
                405                 410                 415
Tyr Phe Ala Phe Asn Pro Lys Leu Cys Val Ser Glu Ile Tyr Arg Met
            420                 425                 430
Glu Glu Val Thr Gly Thr Lys Gly Arg Gln Ser Lys Gly Asp Ile Asn
        435                 440                 445
Thr Arg Asn Asn Gly Glu Arg Ala Ser Cys Glu
    450                 455
```

<210> 127
<211> 146
<212> PRT
<213> Homo sapiens

<400> 127

```
Val Pro Ile Gln Lys Val Gln Asp Asp Thr Lys Thr Leu Ile Lys Thr
1               5                   10                  15
Ile Val Thr Arg Ile Asn Asp Ile Ser His Thr Gln Ser Val Ser Ser
            20                  25                  30
Lys Gln Lys Val Thr Gly Leu Asp Phe Ile Pro Gly Leu His Pro Ile
        35                  40                  45
Leu Thr Leu Ser Lys Met Asp Gln Thr Leu Ala Val Tyr Gln Gln Ile
    50                  55                  60
Leu Thr Ser Met Pro Ser Arg Asn Val Ile Gln Ile Ser Asn Asp Leu
65                  70                  75                  80
Glu Asn Leu Arg Asp Leu Leu His Val Leu Ala Phe Ser Lys Ser Cys
                85                  90                  95
His Leu Pro Trp Ala Ser Gly Leu Glu Thr Leu Asp Ser Leu Gly Gly
            100                 105                 110
Val Leu Glu Ala Ser Gly Tyr Ser Thr Glu Val Val Ala Leu Ser Arg
        115                 120                 125
Leu Gln Gly Ser Leu Gln Asp Met Leu Trp Gln Leu Asp Leu Ser Pro
    130                 135                 140
Gly Cys
145
```

<210> 128
<211> 327
<212> PRT
<213> Homo sapiens

<400> 128

```
Lys Glu Ile Thr Asn Ala Leu Glu Thr Trp Gly Ala Leu Gly Gln Asp
1               5                   10                  15
```

```
Ile Asn Leu Asp Ile Pro Ser Phe Gln Met Ser Asp Asp Ile Asp Asp
            20                  25                  30
Ile Lys Trp Glu Lys Thr Ser Asp Lys Lys Lys Ile Ala Gln Phe Arg
        35                  40                  45
Lys Glu Lys Glu Thr Phe Lys Glu Lys Asp Thr Tyr Lys Leu Phe Lys
    50                  55                  60
Asn Gly Thr Leu Lys Ile Lys His Leu Lys Thr Asp Asp Gln Asp Ile
65                  70                  75                  80
Tyr Lys Val Ser Ile Tyr Asp Thr Lys Gly Lys Asn Val Leu Glu Lys
                85                  90                  95
Ile Phe Asp Leu Lys Ile Gln Glu Arg Val Ser Lys Pro Lys Ile Ser
            100                 105                 110
Trp Thr Cys Ile Asn Thr Thr Leu Thr Cys Glu Val Met Asn Gly Thr
        115                 120                 125
Asp Pro Glu Leu Asn Leu Tyr Gln Asp Gly Lys His Leu Lys Leu Ser
    130                 135                 140
Gln Arg Val Ile Thr His Lys Trp Thr Thr Ser Leu Ser Ala Lys Phe
145                 150                 155                 160
Lys Cys Thr Ala Gly Asn Lys Val Ser Lys Glu Ser Ser Val Glu Pro
                165                 170                 175
Val Ser Cys Pro Glu Lys Gly Leu Asp Ile Tyr Leu Ile Ile Gly Ile
            180                 185                 190
Cys Gly Gly Gly Ser Leu Leu Met Val Phe Val Ala Leu Leu Val Phe
        195                 200                 205
Tyr Ile Thr Lys Arg Lys Lys Gln Arg Ser Arg Arg Asn Asp Glu Glu
    210                 215                 220
Leu Glu Thr Arg Ala His Arg Val Ala Thr Glu Glu Arg Gly Arg Lys
225                 230                 235                 240
Pro Gln Gln Ile Pro Ala Ser Thr Pro Gln Asn Pro Ala Thr Ser Gln
                245                 250                 255
His Pro Pro Pro Pro Pro Gly His Arg Ser Gln Ala Pro Ser His Arg
            260                 265                 270
Pro Pro Pro Pro Gly His Arg Val Gln His Gln Pro Gln Lys Arg Pro
        275                 280                 285
Pro Ala Pro Ser Gly Thr Gln Val His Gln Gln Lys Gly Pro Pro Leu
    290                 295                 300
Pro Arg Pro Arg Val Gln Pro Lys Pro Pro His Gly Ala Ala Glu Asn
305                 310                 315                 320
Ser Leu Ser Pro Ser Ser Asn
                325
```

<210> 129
<211> 433
<212> PRT
<213> Homo sapiens

<400> 129

```
Lys Lys Val Val Leu Gly Lys Lys Gly Asp Thr Val Glu Leu Thr Cys
1               5                   10                  15
Thr Ala Ser Gln Lys Lys Ser Ile Gln Phe His Trp Lys Asn Ser Asn
            20                  25                  30
Gln Ile Lys Ile Leu Gly Asn Gln Gly Ser Phe Leu Thr Lys Gly Pro
        35                  40                  45
Ser Lys Leu Asn Asp Arg Ala Asp Ser Arg Arg Ser Leu Trp Asp Gln
    50                  55                  60
Gly Asn Phe Pro Leu Ile Ile Lys Asn Leu Lys Ile Glu Asp Ser Asp
65                  70                  75                  80
Thr Tyr Ile Cys Glu Val Glu Asp Gln Lys Glu Glu Val Gln Leu Leu
                85                  90                  95
Val Phe Gly Leu Thr Ala Asn Ser Asp Thr His Leu Leu Gln Gly Gln
            100                 105                 110
Ser Leu Thr Leu Thr Leu Glu Ser Pro Pro Gly Ser Ser Pro Ser Val
```

```
        115                 120                 125
Gln Cys Arg Ser Pro Arg Gly Lys Asn Ile Gln Gly Gly Lys Thr Leu
    130                 135                 140
Ser Val Ser Gln Leu Glu Leu Gln Asp Ser Gly Thr Trp Thr Cys Thr
145                 150                 155                 160
Val Leu Gln Asn Gln Lys Lys Val Glu Phe Lys Ile Asp Ile Val Val
                165                 170                 175
Leu Ala Phe Gln Lys Ala Ser Ser Ile Val Tyr Lys Lys Glu Gly Glu
            180                 185                 190
Gln Val Glu Phe Ser Phe Pro Leu Ala Phe Thr Val Glu Lys Leu Thr
        195                 200                 205
Gly Ser Gly Glu Leu Trp Trp Gln Ala Glu Arg Ala Ser Ser Ser Lys
    210                 215                 220
Ser Trp Ile Thr Phe Asp Leu Lys Asn Lys Glu Val Ser Val Lys Arg
225                 230                 235                 240
Val Thr Gln Asp Pro Lys Leu Gln Met Gly Lys Lys Leu Pro Leu His
                245                 250                 255
Leu Thr Leu Pro Gln Ala Leu Pro Gln Tyr Ala Gly Ser Gly Asn Leu
            260                 265                 270
Thr Leu Ala Leu Glu Ala Lys Thr Gly Lys Leu His Gln Glu Val Asn
        275                 280                 285
Leu Val Val Met Arg Ala Thr Gln Leu Gln Lys Asn Leu Thr Cys Glu
    290                 295                 300
Val Trp Gly Pro Thr Ser Pro Lys Leu Met Leu Ser Leu Lys Leu Glu
305                 310                 315                 320
Asn Lys Glu Ala Lys Val Ser Lys Arg Glu Lys Ala Val Trp Val Leu
                325                 330                 335
Asn Pro Glu Ala Gly Met Trp Gln Cys Leu Leu Ser Asp Ser Gly Gln
            340                 345                 350
Val Leu Leu Glu Ser Asn Ile Lys Val Leu Pro Thr Trp Ser Thr Pro
        355                 360                 365
Val Gln Pro Met Ala Leu Ile Val Leu Gly Gly Val Ala Gly Leu Leu
    370                 375                 380
Leu Phe Ile Gly Leu Gly Ile Phe Phe Cys Val Arg Cys Arg His Arg
385                 390                 395                 400
Arg Arg Gln Ala Glu Arg Met Ser Gln Ile Lys Arg Leu Leu Ser Glu
                405                 410                 415
Lys Lys Thr Cys Gln Cys Pro His Arg Phe Gln Lys Thr Cys Ser Pro
            420                 425                 430
Ile
```

<210> 130
<211> 1145
<212> PRT
<213> Homo sapiens

<400> 130

```
Tyr Asn Leu Asp Val Arg Gly Ala Arg Ser Phe Ser Pro Pro Arg Ala
1               5                   10                  15
Gly Arg His Phe Gly Tyr Arg Val Leu Gln Val Gly Asn Gly Val Ile
            20                  25                  30
Val Gly Ala Pro Gly Glu Gly Asn Ser Thr Gly Ser Leu Tyr Gln Cys
        35                  40                  45
Gln Ser Gly Thr Gly His Cys Leu Pro Val Thr Leu Arg Gly Ser Asn
    50                  55                  60
Tyr Thr Ser Lys Tyr Leu Gly Met Thr Leu Ala Thr Asp Pro Thr Asp
65                  70                  75                  80
Gly Ser Ile Leu Ala Cys Asp Pro Gly Leu Ser Arg Thr Cys Asp Gln
                85                  90                  95
Asn Thr Tyr Leu Ser Gly Leu Cys Tyr Leu Phe Arg Gln Asn Leu Gln
            100                 105                 110
Gly Pro Met Leu Gln Gly Arg Pro Gly Phe Gln Glu Cys Ile Lys Gly
```

```
        115                 120                 125
Asn Val Asp Leu Val Phe Leu Phe Asp Gly Ser Met Ser Leu Gln Pro
    130                 135                 140
Asp Glu Phe Gln Lys Ile Leu Asp Phe Met Lys Asp Val Met Lys Lys
145                 150                 155                 160
Leu Ser Asn Thr Ser Tyr Gln Phe Ala Ala Val Gln Phe Ser Thr Ser
                165                 170                 175
Tyr Lys Thr Glu Phe Asp Phe Ser Asp Tyr Val Lys Arg Lys Asp Pro
            180                 185                 190
Asp Ala Leu Leu Lys His Val Lys His Met Leu Leu Leu Thr Asn Thr
        195                 200                 205
Phe Gly Ala Ile Asn Tyr Val Ala Thr Glu Val Phe Arg Glu Glu Leu
    210                 215                 220
Gly Ala Arg Pro Asp Ala Thr Lys Val Leu Ile Ile Ile Thr Asp Gly
225                 230                 235                 240
Glu Ala Thr Asp Ser Gly Asn Ile Asp Ala Ala Lys Asp Ile Ile Arg
                245                 250                 255
Tyr Ile Ile Gly Ile Gly Lys His Phe Gln Thr Lys Glu Ser Gln Glu
            260                 265                 270
Thr Leu His Lys Phe Ala Ser Lys Pro Ala Ser Glu Phe Val Lys Ile
        275                 280                 285
Leu Asp Thr Phe Glu Lys Leu Lys Asp Leu Phe Thr Glu Leu Gln Lys
    290                 295                 300
Lys Ile Tyr Val Ile Glu Gly Thr Ser Lys Gln Asp Leu Thr Ser Phe
305                 310                 315                 320
Asn Met Glu Leu Ser Ser Ser Gly Ile Ser Ala Asp Leu Ser Arg Gly
                325                 330                 335
His Ala Val Val Gly Ala Val Gly Ala Lys Asp Trp Ala Gly Gly Phe
            340                 345                 350
Leu Asp Leu Lys Ala Asp Leu Gln Asp Asp Thr Phe Ile Gly Asn Glu
        355                 360                 365
Pro Leu Thr Pro Glu Val Arg Ala Gly Tyr Leu Gly Tyr Thr Val Thr
    370                 375                 380
Trp Leu Pro Ser Arg Gln Lys Thr Ser Leu Leu Ala Ser Gly Ala Pro
385                 390                 395                 400
Arg Tyr Gln His Met Gly Arg Val Leu Leu Phe Gln Glu Pro Gln Gly
                405                 410                 415
Gly Gly His Trp Ser Gln Val Gln Thr Ile His Gly Thr Gln Ile Gly
            420                 425                 430
Ser Tyr Phe Gly Gly Glu Leu Cys Gly Val Asp Val Asp Gln Asp Gly
        435                 440                 445
Glu Thr Glu Leu Leu Leu Ile Gly Ala Pro Leu Phe Tyr Gly Glu Gln
    450                 455                 460
Arg Gly Gly Arg Val Phe Ile Tyr Gln Arg Arg Gln Leu Gly Phe Glu
465                 470                 475                 480
Glu Val Ser Glu Leu Gln Gly Asp Pro Gly Tyr Pro Leu Gly Arg Phe
                485                 490                 495
Gly Glu Ala Ile Thr Ala Leu Thr Asp Ile Asn Gly Asp Gly Leu Val
            500                 505                 510
Asp Val Ala Val Gly Ala Pro Leu Glu Glu Gln Gly Ala Val Tyr Ile
        515                 520                 525
Phe Asn Gly Arg His Gly Gly Leu Ser Pro Gln Pro Ser Gln Arg Ile
    530                 535                 540
Glu Gly Thr Gln Val Leu Ser Gly Ile Gln Trp Phe Gly Arg Ser Ile
545                 550                 555                 560
His Gly Val Lys Asp Leu Glu Gly Asp Gly Leu Ala Asp Val Ala Val
                565                 570                 575
Gly Ala Glu Ser Gln Met Ile Val Leu Ser Ser Arg Pro Val Val Asp
            580                 585                 590
Met Val Thr Leu Met Ser Phe Ser Pro Ala Glu Ile Pro Val His Glu
        595                 600                 605
Val Glu Cys Ser Tyr Ser Thr Ser Asn Lys Met Lys Glu Gly Val Asn
    610                 615                 620
```

```
Ile Thr Ile Cys Phe Gln Ile Lys Ser Leu Tyr Pro Gln Phe Gln Gly
625                 630                 635                 640
Arg Leu Val Ala Asn Leu Thr Tyr Thr Leu Gln Leu Asp Gly His Arg
                645                 650                 655
Thr Arg Arg Arg Gly Leu Phe Pro Gly Gly Arg His Glu Leu Arg Arg
            660                 665                 670
Asn Ile Ala Val Thr Thr Ser Met Ser Cys Thr Asp Phe Ser Phe His
        675                 680                 685
Phe Pro Val Cys Val Gln Asp Leu Ile Ser Pro Ile Asn Val Ser Leu
    690                 695                 700
Asn Phe Ser Leu Trp Glu Glu Glu Gly Thr Pro Arg Asp Gln Arg Ala
705                 710                 715                 720
Gln Gly Lys Asp Ile Pro Pro Ile Leu Arg Pro Ser Leu His Ser Glu
                725                 730                 735
Thr Trp Glu Ile Pro Phe Glu Lys Asn Cys Gly Glu Asp Lys Lys Cys
            740                 745                 750
Glu Ala Asn Leu Arg Val Ser Phe Ser Pro Ala Arg Ser Arg Ala Leu
        755                 760                 765
Arg Leu Thr Ala Phe Ala Ser Leu Ser Val Glu Leu Ser Leu Ser Asn
    770                 775                 780
Leu Glu Glu Asp Ala Tyr Trp Val Gln Leu Asp Leu His Phe Pro Pro
785                 790                 795                 800
Gly Leu Ser Phe Arg Lys Val Glu Met Leu Lys Pro His Ser Gln Ile
                805                 810                 815
Pro Val Ser Cys Glu Glu Leu Pro Glu Glu Ser Arg Leu Leu Ser Arg
            820                 825                 830
Ala Leu Ser Cys Asn Val Ser Ser Pro Ile Phe Lys Ala Gly His Ser
        835                 840                 845
Val Ala Leu Gln Met Met Phe Asn Thr Leu Val Asn Ser Ser Trp Gly
    850                 855                 860
Asp Ser Val Glu Leu His Ala Asn Val Thr Cys Asn Asn Glu Asp Ser
865                 870                 875                 880
Asp Leu Leu Glu Asp Asn Ser Ala Thr Thr Ile Ile Pro Ile Leu Tyr
                885                 890                 895
Pro Ile Asn Ile Leu Ile Gln Asp Gln Glu Asp Ser Thr Leu Tyr Val
            900                 905                 910
Ser Phe Thr Pro Lys Gly Pro Lys Ile His Gln Val Lys His Met Tyr
        915                 920                 925
Gln Val Arg Ile Gln Pro Ser Ile His Asp His Asn Ile Pro Thr Leu
    930                 935                 940
Glu Ala Val Val Gly Val Pro Gln Pro Pro Ser Glu Gly Pro Ile Thr
945                 950                 955                 960
His Gln Trp Ser Val Gln Met Glu Pro Pro Val Pro Cys His Tyr Glu
                965                 970                 975
Asp Leu Glu Arg Leu Pro Asp Ala Ala Glu Pro Cys Leu Pro Gly Ala
            980                 985                 990
Leu Phe Arg Cys Pro Val Val Phe  Arg Gln Glu Ile Leu  Val Gln Val
        995                1000                1005
Ile Gly  Thr Leu Glu Leu Val  Gly Glu Ile Glu Ala  Ser Ser Met
    1010                1015                1020
Phe Ser  Leu Cys Ser Ser Leu  Ser Ile Ser Phe Asn  Ser Ser Lys
    1025                1030                1035
His Phe  His Leu Tyr Gly Ser  Asn Ala Ser Leu Ala  Gln Val Val
    1040                1045                1050
Met Lys  Val Asp Val Val Tyr  Glu Lys Gln Met Leu  Tyr Leu Tyr
    1055                1060                1065
Val Leu  Ser Gly Ile Gly Gly  Leu Leu Leu Leu Leu  Leu Ile Phe
    1070                1075                1080
Ile Val  Leu Tyr Lys Val Gly  Phe Phe Lys Arg Asn  Leu Lys Glu
    1085                1090                1095
Lys Met  Glu Ala Gly Arg Gly  Val Pro Asn Gly Ile  Pro Ala Glu
    1100                1105                1110
Asp Ser  Glu Gln Leu Ala Ser  Gly Gln Glu Ala Gly  Asp Pro Gly
```

1115 1120 1125
Cys Leu Lys Pro Leu His Glu Lys Asp Ser Glu Ser Gly Gly Gly
1130 1135 1140
Lys Asp
1145

<210> 131
<211> 660
<212> PRT
<213> Homo sapiens

<400> 131
Met Glu Ala Leu Met Ala Arg Gly Ala Leu Thr Gly Pro Leu Arg Ala
1 5 10 15
Leu Cys Leu Leu Gly Cys Leu Leu Ser His Ala Ala Ala Ala Pro Ser
20 25 30
Pro Ile Ile Lys Phe Pro Gly Asp Val Ala Pro Lys Thr Asp Lys Glu
35 40 45
Leu Ala Val Gln Tyr Leu Asn Thr Phe Tyr Gly Cys Pro Lys Glu Ser
50 55 60
Cys Asn Leu Phe Val Leu Lys Asp Thr Leu Lys Lys Met Gln Lys Phe
65 70 75 80
Phe Gly Leu Pro Gln Thr Gly Asp Leu Asp Gln Asn Thr Ile Glu Thr
85 90 95
Met Arg Lys Pro Arg Cys Gly Asn Pro Asp Val Ala Asn Tyr Asn Phe
100 105 110
Phe Pro Arg Lys Pro Lys Trp Asp Lys Asn Gln Ile Thr Tyr Arg Ile
115 120 125
Ile Gly Tyr Thr Pro Asp Leu Asp Pro Glu Thr Val Asp Asp Ala Phe
130 135 140
Ala Arg Ala Phe Gln Val Trp Ser Asp Val Thr Pro Leu Arg Phe Ser
145 150 155 160
Arg Ile His Asp Gly Glu Ala Asp Ile Met Ile Asn Phe Gly Arg Trp
165 170 175
Glu His Gly Asp Gly Tyr Pro Phe Asp Gly Lys Asp Gly Leu Leu Ala
180 185 190
His Ala Phe Ala Pro Gly Thr Gly Val Gly Gly Asp Ser His Phe Asp
195 200 205
Asp Asp Glu Leu Trp Thr Leu Gly Glu Gly Gln Val Val Arg Val Lys
210 215 220
Tyr Gly Asn Ala Asp Gly Glu Tyr Cys Lys Phe Pro Phe Leu Phe Asn
225 230 235 240
Gly Lys Glu Tyr Asn Ser Cys Thr Asp Thr Gly Arg Ser Asp Gly Phe
245 250 255
Leu Trp Cys Ser Thr Thr Tyr Asn Phe Glu Lys Asp Gly Lys Tyr Gly
260 265 270
Phe Cys Pro His Glu Ala Leu Phe Thr Met Gly Gly Asn Ala Glu Gly
275 280 285
Gln Pro Cys Lys Phe Pro Phe Arg Phe Gln Gly Thr Ser Tyr Asp Ser
290 295 300
Cys Thr Thr Glu Gly Arg Thr Asp Gly Tyr Arg Trp Cys Gly Thr Thr
305 310 315 320
Glu Asp Tyr Asp Arg Asp Lys Lys Tyr Gly Phe Cys Pro Glu Thr Ala
325 330 335
Met Ser Thr Val Gly Gly Asn Ser Glu Gly Ala Pro Cys Val Phe Pro
340 345 350
Phe Thr Phe Leu Gly Asn Lys Tyr Glu Ser Cys Thr Ser Ala Gly Arg
355 360 365
Ser Asp Gly Lys Met Trp Cys Ala Thr Thr Ala Asn Tyr Asp Asp Asp
370 375 380
Arg Lys Trp Gly Phe Cys Pro Asp Gln Gly Tyr Ser Leu Phe Leu Val
385 390 395 400

```
Ala Ala His Glu Phe Gly His Ala Met Gly Leu Glu His Ser Gln Asp
                405                 410                 415
Pro Gly Ala Leu Met Ala Pro Ile Tyr Thr Tyr Thr Lys Asn Phe Arg
            420                 425                 430
Leu Ser Gln Asp Asp Ile Lys Gly Ile Gln Glu Leu Tyr Gly Ala Ser
        435                 440                 445
Pro Asp Ile Asp Leu Gly Thr Gly Pro Thr Pro Thr Leu Gly Pro Val
    450                 455                 460
Thr Pro Glu Ile Cys Lys Gln Asp Ile Val Phe Asp Gly Ile Ala Gln
465                 470                 475                 480
Ile Arg Gly Glu Ile Phe Phe Phe Lys Asp Arg Phe Ile Trp Arg Thr
                485                 490                 495
Val Thr Pro Arg Asp Lys Pro Met Gly Pro Leu Leu Val Ala Thr Phe
            500                 505                 510
Trp Pro Glu Leu Pro Glu Lys Ile Asp Ala Val Tyr Glu Ala Pro Gln
        515                 520                 525
Glu Glu Lys Ala Val Phe Phe Ala Gly Asn Glu Tyr Trp Ile Tyr Ser
    530                 535                 540
Ala Ser Thr Leu Glu Arg Gly Tyr Pro Lys Pro Leu Thr Ser Leu Gly
545                 550                 555                 560
Leu Pro Pro Asp Val Gln Arg Val Asp Ala Ala Phe Asn Trp Ser Lys
                565                 570                 575
Asn Lys Lys Thr Tyr Ile Phe Ala Gly Asp Lys Phe Trp Arg Tyr Asn
            580                 585                 590
Glu Val Lys Lys Lys Met Asp Pro Gly Phe Pro Lys Leu Ile Ala Asp
        595                 600                 605
Ala Trp Asn Ala Ile Pro Asp Asn Leu Asp Ala Val Val Asp Leu Gln
    610                 615                 620
Gly Gly Gly His Ser Tyr Phe Phe Lys Gly Ala Tyr Tyr Leu Lys Leu
625                 630                 635                 640
Glu Asn Gln Ser Leu Lys Ser Val Lys Phe Gly Ser Ile Lys Ser Asp
                645                 650                 655
Trp Leu Gly Cys
            660
```

<210> 132
<211> 707
<212> PRT
<213> Homo sapiens

<400> 132

```
Met Ser Leu Trp Gln Pro Leu Val Leu Val Leu Leu Val Leu Gly Cys
1               5                   10                  15
Cys Phe Ala Ala Pro Arg Gln Arg Gln Ser Thr Leu Val Leu Phe Pro
            20                  25                  30
Gly Asp Leu Arg Thr Asn Leu Thr Asp Arg Gln Leu Ala Glu Glu Tyr
        35                  40                  45
Leu Tyr Arg Tyr Gly Tyr Thr Arg Val Ala Glu Met Arg Gly Glu Ser
    50                  55                  60
Lys Ser Leu Gly Pro Ala Leu Leu Leu Leu Gln Lys Gln Leu Ser Leu
65                  70                  75                  80
Pro Glu Thr Gly Glu Leu Asp Ser Ala Thr Leu Lys Ala Met Arg Thr
                85                  90                  95
Pro Arg Cys Gly Val Pro Asp Leu Gly Arg Phe Gln Thr Phe Glu Gly
            100                 105                 110
Asp Leu Lys Trp His His His Asn Ile Thr Tyr Trp Ile Gln Asn Tyr
        115                 120                 125
Ser Glu Asp Leu Pro Arg Ala Val Ile Asp Asp Ala Phe Ala Arg Ala
    130                 135                 140
Phe Ala Leu Trp Ser Ala Val Thr Pro Leu Thr Phe Thr Arg Val Tyr
145                 150                 155                 160
Ser Arg Asp Ala Asp Ile Val Ile Gln Phe Gly Val Ala Glu His Gly
```

```
                 165                 170                 175
Asp Gly Tyr Pro Phe Asp Gly Lys Asp Gly Leu Leu Ala His Ala Phe
            180                 185                 190
Pro Pro Gly Pro Gly Ile Gln Gly Asp Ala His Phe Asp Asp Asp Glu
        195                 200                 205
Leu Trp Ser Leu Gly Lys Gly Val Val Val Pro Thr Arg Phe Gly Asn
    210                 215                 220
Ala Asp Gly Ala Ala Cys His Phe Pro Phe Ile Phe Glu Gly Arg Ser
225                 230                 235                 240
Tyr Ser Ala Cys Thr Thr Asp Gly Arg Ser Asp Gly Leu Pro Trp Cys
                245                 250                 255
Ser Thr Thr Ala Asn Tyr Asp Thr Asp Asp Arg Phe Gly Phe Cys Pro
            260                 265                 270
Ser Glu Arg Leu Tyr Thr Arg Asp Gly Asn Ala Asp Gly Lys Pro Cys
        275                 280                 285
Gln Phe Pro Phe Ile Phe Gln Gly Gln Ser Tyr Ser Ala Cys Thr Thr
    290                 295                 300
Asp Gly Arg Ser Asp Gly Tyr Arg Trp Cys Ala Thr Thr Ala Asn Tyr
305                 310                 315                 320
Asp Arg Asp Lys Leu Phe Gly Phe Cys Pro Thr Arg Ala Asp Ser Thr
                325                 330                 335
Val Met Gly Gly Asn Ser Ala Gly Glu Leu Cys Val Phe Pro Phe Thr
            340                 345                 350
Phe Leu Gly Lys Glu Tyr Ser Thr Cys Thr Ser Glu Gly Arg Gly Asp
        355                 360                 365
Gly Arg Leu Trp Cys Ala Thr Thr Ser Asn Phe Asp Ser Asp Lys Lys
    370                 375                 380
Trp Gly Phe Cys Pro Asp Gln Gly Tyr Ser Leu Phe Leu Val Ala Ala
385                 390                 395                 400
His Glu Phe Gly His Ala Leu Gly Leu Asp His Ser Ser Val Pro Glu
                405                 410                 415
Ala Leu Met Tyr Pro Met Tyr Arg Phe Thr Glu Gly Pro Pro Leu His
            420                 425                 430
Lys Asp Asp Val Asn Gly Ile Arg His Leu Tyr Gly Pro Arg Pro Glu
        435                 440                 445
Pro Glu Pro Arg Pro Pro Thr Thr Thr Thr Pro Gln Pro Thr Ala Pro
    450                 455                 460
Pro Thr Val Cys Pro Thr Gly Pro Pro Thr Val His Pro Ser Glu Arg
465                 470                 475                 480
Pro Thr Ala Gly Pro Thr Gly Pro Pro Ser Ala Gly Pro Thr Gly Pro
                485                 490                 495
Pro Thr Ala Gly Pro Ser Thr Ala Thr Thr Val Pro Leu Ser Pro Val
            500                 505                 510
Asp Asp Ala Cys Asn Val Asn Ile Phe Asp Ala Ile Ala Glu Ile Gly
        515                 520                 525
Asn Gln Leu Tyr Leu Phe Lys Asp Gly Lys Tyr Trp Arg Phe Ser Glu
    530                 535                 540
Gly Arg Gly Ser Arg Pro Gln Gly Pro Phe Leu Ile Ala Asp Lys Trp
545                 550                 555                 560
Pro Ala Leu Pro Arg Lys Leu Asp Ser Val Phe Glu Glu Pro Leu Ser
                565                 570                 575
Lys Lys Leu Phe Phe Phe Ser Gly Arg Gln Val Trp Val Tyr Thr Gly
            580                 585                 590
Ala Ser Val Leu Gly Pro Arg Arg Leu Asp Lys Leu Gly Leu Gly Ala
        595                 600                 605
Asp Val Ala Gln Val Thr Gly Ala Leu Arg Ser Gly Arg Gly Lys Met
    610                 615                 620
Leu Leu Phe Ser Gly Arg Arg Leu Trp Arg Phe Asp Val Lys Ala Gln
625                 630                 635                 640
Met Val Asp Pro Arg Ser Ala Ser Glu Val Asp Arg Met Phe Pro Gly
                645                 650                 655
Val Pro Leu Asp Thr His Asp Val Phe Gln Tyr Arg Glu Lys Ala Tyr
            660                 665                 670
```

Phe Cys Gln Asp Arg Phe Tyr Trp Arg Val Ser Ser Arg Ser Glu Leu
675 680 685
Asn Gln Val Asp Gln Val Gly Tyr Val Thr Tyr Asp Ile Leu Gln Cys
690 695 700
Pro Glu Asp
705

<210> 133
<211> 115
<212> PRT
<213> Homo sapiens

<400> 133
Ile Pro Thr Glu Ile Pro Thr Ser Ala Leu Val Lys Glu Thr Leu Ala
1 5 10 15
Leu Leu Ser Thr His Arg Thr Leu Leu Ile Ala Asn Glu Thr Leu Arg
20 25 30
Ile Pro Val Pro Val His Lys Asn His Gln Leu Cys Thr Glu Glu Ile
35 40 45
Phe Gln Gly Ile Gly Thr Leu Glu Ser Gln Thr Val Gln Gly Gly Thr
50 55 60
Val Glu Arg Leu Phe Lys Asn Leu Ser Leu Ile Lys Lys Tyr Ile Asp
65 70 75 80
Gly Gln Lys Lys Lys Cys Gly Glu Glu Arg Arg Arg Val Asn Gln Phe
85 90 95
Leu Asp Tyr Leu Gln Glu Phe Leu Gly Val Met Asn Thr Glu Trp Ile
100 105 110
Ile Glu Ser
115

<210> 134
<211> 185
<212> PRT
<213> Homo sapiens

<400> 134
Ala Pro Val Pro Pro Gly Glu Asp Ser Lys Asp Val Ala Ala Pro His
1 5 10 15
Arg Gln Pro Leu Thr Ser Ser Glu Arg Ile Asp Lys Gln Ile Arg Tyr
20 25 30
Ile Leu Asp Gly Ile Ser Ala Leu Arg Lys Glu Thr Cys Asn Lys Ser
35 40 45
Asn Met Cys Glu Ser Ser Lys Glu Ala Leu Ala Glu Asn Asn Leu Asn
50 55 60
Leu Pro Lys Met Ala Glu Lys Asp Gly Cys Phe Gln Ser Gly Phe Asn
65 70 75 80
Glu Glu Thr Cys Leu Val Lys Ile Ile Thr Gly Leu Leu Glu Phe Glu
85 90 95
Val Tyr Leu Glu Tyr Leu Gln Asn Arg Phe Glu Ser Ser Glu Glu Gln
100 105 110
Ala Arg Ala Val Gln Met Ser Thr Lys Val Leu Ile Gln Phe Leu Gln
115 120 125
Lys Lys Ala Lys Asn Leu Asp Ala Ile Thr Thr Pro Asp Pro Thr Thr
130 135 140
Asn Ala Ser Leu Leu Thr Lys Leu Gln Ala Gln Asn Gln Trp Leu Gln
145 150 155 160
Asp Met Thr Thr His Leu Ile Leu Arg Ser Phe Lys Glu Phe Leu Gln
165 170 175
Ser Ser Leu Arg Ala Leu Arg Gln Met
180 185

<210> 135
<211> 160
<212> PRT
<213> Homo sapiens

<400> 135
Ser Pro Gly Gln Gly Thr Gln Ser Glu Asn Ser Cys Thr His Phe Pro
1 5 10 15
Gly Asn Leu Pro Asn Met Leu Arg Asp Leu Arg Asp Ala Phe Ser Arg
20 25 30
Val Lys Thr Phe Phe Gln Met Lys Asp Gln Leu Asp Asn Leu Leu Leu
35 40 45
Lys Glu Ser Leu Leu Glu Asp Phe Lys Gly Tyr Leu Gly Cys Gln Ala
50 55 60
Leu Ser Glu Met Ile Gln Phe Tyr Leu Glu Glu Val Met Pro Gln Ala
65 70 75 80
Glu Asn Gln Asp Pro Asp Ile Lys Ala His Val Asn Ser Leu Gly Glu
85 90 95
Asn Leu Lys Thr Leu Arg Leu Arg Leu Arg Arg Cys His Arg Phe Leu
100 105 110
Pro Cys Glu Asn Lys Ser Lys Ala Val Glu Gln Val Lys Asn Ala Phe
115 120 125
Asn Lys Leu Gln Glu Lys Gly Ile Tyr Lys Ala Met Ser Glu Phe Asp
130 135 140
Ile Phe Ile Asn Tyr Ile Glu Ala Tyr Met Thr Met Lys Ile Arg Asn
145 150 155 160

<210> 136
<211> 472
<212> PRT
<213> Homo sapiens

<400> 136
Glu Met Gly Thr Ala Asp Leu Gly Pro Ser Ser Val Pro Thr Pro Thr
1 5 10 15
Asn Val Thr Ile Glu Ser Tyr Asn Met Asn Pro Ile Val Tyr Trp Glu
20 25 30
Tyr Gln Ile Met Pro Gln Val Pro Val Phe Thr Val Glu Val Lys Asn
35 40 45
Tyr Gly Val Lys Asn Ser Glu Trp Ile Asp Ala Cys Ile Asn Ile Ser
50 55 60
His His Tyr Cys Asn Ile Ser Asp His Val Gly Asp Pro Ser Asn Ser
65 70 75 80
Leu Trp Val Arg Val Lys Ala Arg Val Gly Gln Lys Glu Ser Ala Tyr
85 90 95
Ala Lys Ser Glu Glu Phe Ala Val Cys Arg Asp Gly Lys Ile Gly Pro
100 105 110
Pro Lys Leu Asp Ile Arg Lys Glu Glu Lys Gln Ile Met Ile Asp Ile
115 120 125
Phe His Pro Ser Val Phe Val Asn Gly Asp Glu Gln Glu Val Asp Tyr
130 135 140
Asp Pro Glu Thr Thr Cys Tyr Ile Arg Val Tyr Asn Val Tyr Val Arg
145 150 155 160
Met Asn Gly Ser Glu Ile Gln Tyr Lys Ile Leu Thr Gln Lys Glu Asp
165 170 175
Asp Cys Asp Glu Ile Gln Cys Gln Leu Ala Ile Pro Val Ser Ser Leu
180 185 190
Asn Ser Gln Tyr Cys Val Ser Ala Glu Gly Val Leu His Val Trp Gly
195 200 205
Val Thr Thr Glu Lys Ser Lys Glu Val Cys Ile Thr Ile Phe Asn Ser
210 215 220
Ser Ile Lys Gly Ser Leu Trp Ile Pro Val Val Ala Ala Leu Leu Leu

225 230 235 240
Phe Leu Val Leu Ser Leu Val Phe Ile Cys Phe Tyr Ile Lys Lys Ile
245 250 255
Asn Pro Leu Lys Glu Lys Ser Ile Ile Leu Pro Lys Ser Leu Ile Ser
260 265 270
Val Val Arg Ser Ala Thr Leu Glu Thr Lys Pro Glu Ser Lys Tyr Val
275 280 285
Ser Leu Ile Thr Ser Tyr Gln Pro Phe Ser Leu Glu Lys Glu Val Val
290 295 300
Cys Glu Glu Pro Leu Ser Pro Ala Thr Val Pro Gly Met His Thr Glu
305 310 315 320
Asp Asn Pro Gly Lys Val Glu His Thr Glu Glu Leu Ser Ser Ile Thr
325 330 335
Glu Val Val Thr Thr Glu Glu Asn Ile Pro Asp Val Val Pro Gly Ser
340 345 350
His Leu Thr Pro Ile Glu Arg Glu Ser Ser Ser Pro Leu Ser Ser Asn
355 360 365
Gln Ser Glu Pro Gly Ser Ile Ala Leu Asn Ser Tyr His Ser Arg Asn
370 375 380
Cys Ser Glu Ser Asp His Ser Arg Asn Gly Phe Asp Thr Asp Ser Ser
385 390 395 400
Cys Leu Glu Ser His Ser Ser Leu Ser Asp Ser Glu Phe Pro Pro Asn
405 410 415
Asn Lys Gly Glu Ile Lys Thr Glu Gly Gln Glu Leu Ile Thr Val Ile
420 425 430
Lys Ala Pro Thr Ser Phe Gly Tyr Asp Lys Pro His Val Leu Val Asp
435 440 445
Leu Leu Val Asp Asp Ser Gly Lys Glu Ser Leu Ile Gly Tyr Arg Pro
450 455 460
Thr Glu Asp Ser Lys Glu Phe Ser
465 470

<210> 137
<211> 143
<212> PRT
<213> Homo sapiens

<400> 137
Gln Asp Pro Tyr Val Lys Glu Ala Glu Asn Leu Lys Lys Tyr Phe Asn
1 5 10 15
Ala Gly His Ser Asp Val Ala Asp Asn Gly Thr Leu Phe Leu Gly Ile
20 25 30
Leu Lys Asn Trp Lys Glu Glu Ser Asp Arg Lys Ile Met Gln Ser Gln
35 40 45
Ile Val Ser Phe Tyr Phe Lys Leu Phe Lys Asn Phe Lys Asp Asp Gln
50 55 60
Ser Ile Gln Lys Ser Val Glu Thr Ile Lys Glu Asp Met Asn Val Lys
65 70 75 80
Phe Phe Asn Ser Asn Lys Lys Lys Arg Asp Asp Phe Glu Lys Leu Thr
85 90 95
Asn Tyr Ser Val Thr Asp Leu Asn Val Gln Arg Lys Ala Ile His Glu
100 105 110
Leu Ile Gln Val Met Ala Glu Leu Ser Pro Ala Ala Lys Thr Gly Lys
115 120 125
Arg Lys Arg Ser Gln Met Leu Phe Arg Gly Arg Arg Ala Ser Gln
130 135 140

<210> 138
<211> 143
<212> PRT

<213> Homo sapiens

<400> 138

```
Met Glu Ser Pro Ser Ala Pro Pro His Arg Trp Cys Ile Pro Trp Gln
1               5                  10                  15
Arg Leu Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro Pro Thr
            20                  25                  30
Thr Ala Lys Leu Thr Ile Glu Ser Thr Pro Phe Asn Val Ala Glu Gly
        35                  40                  45
Lys Glu Val Leu Leu Leu Val His Asn Leu Pro Gln His Leu Phe Gly
    50                  55                  60
Tyr Ser Trp Tyr Lys Gly Glu Arg Val Asp Gly Asn Arg Gln Ile Ile
65                  70                  75                  80
Gly Tyr Val Ile Gly Thr Gln Gln Ala Thr Pro Gly Pro Ala Tyr Ser
                85                  90                  95
Gly Arg Glu Ile Ile Tyr Pro Asn Ala Ser Leu Leu Ile Gln Asn Ile
            100                 105                 110
Ile Gln Asn Asp Thr Gly Phe Tyr Thr Leu His Val Ile Lys Ser Asp
        115                 120                 125
Leu Val Asn Glu Glu Ala Thr Gly Gln Phe Arg Val Tyr Arg Glu
    130                 135                 140
```

<210> 139
<211> 440
<212> PRT
<213> Homo sapiens

<400> 139

```
Glu Met Val Asp Asn Leu Arg Gly Lys Ser Gly Gln Gly Tyr Tyr Val
1               5                  10                  15
Glu Met Thr Val Gly Ser Pro Pro Gln Thr Leu Asn Ile Leu Val Asp
            20                  25                  30
Thr Gly Ser Ser Asn Phe Ala Val Gly Ala Ala Pro His Pro Phe Leu
        35                  40                  45
His Arg Tyr Tyr Gln Arg Gln Leu Ser Ser Thr Tyr Arg Asp Leu Arg
    50                  55                  60
Lys Gly Val Tyr Val Pro Tyr Thr Gln Gly Lys Trp Glu Gly Glu Leu
65                  70                  75                  80
Gly Thr Asp Leu Val Ser Ile Pro His Gly Pro Asn Val Thr Val Arg
                85                  90                  95
Ala Asn Ile Ala Ala Ile Thr Glu Ser Asp Lys Phe Phe Ile Asn Gly
            100                 105                 110
Ser Asn Trp Glu Gly Ile Leu Gly Leu Ala Tyr Ala Glu Ile Ala Arg
        115                 120                 125
Pro Asp Asp Ser Leu Glu Pro Phe Phe Asp Ser Leu Val Lys Gln Thr
    130                 135                 140
His Val Pro Asn Leu Phe Ser Leu Gln Leu Cys Gly Ala Gly Phe Pro
145                 150                 155                 160
Leu Asn Gln Ser Glu Val Leu Ala Ser Val Gly Gly Ser Met Ile Ile
                165                 170                 175
Gly Gly Ile Asp His Ser Leu Tyr Thr Gly Ser Leu Trp Tyr Thr Pro
            180                 185                 190
Ile Arg Arg Glu Trp Tyr Tyr Glu Val Ile Ile Val Arg Val Glu Ile
        195                 200                 205
Asn Gly Gln Asp Leu Lys Met Asp Cys Lys Glu Tyr Asn Tyr Asp Lys
    210                 215                 220
Ser Ile Val Asp Ser Gly Thr Thr Asn Leu Arg Leu Pro Lys Lys Val
225                 230                 235                 240
Phe Glu Ala Ala Val Lys Ser Ile Lys Ala Ala Ser Ser Thr Glu Lys
                245                 250                 255
Phe Pro Asp Gly Phe Trp Leu Gly Glu Gln Leu Val Cys Trp Gln Ala
            260                 265                 270
```

```
Gly Thr Thr Pro Trp Asn Ile Phe Pro Val Ile Ser Leu Tyr Leu Met
        275                 280                 285
Gly Glu Val Thr Asn Gln Ser Phe Arg Ile Thr Ile Leu Pro Gln Gln
    290                 295                 300
Tyr Leu Arg Pro Val Glu Asp Val Ala Thr Ser Gln Asp Asp Cys Tyr
305                 310                 315                 320
Lys Phe Ala Ile Ser Gln Ser Ser Thr Gly Thr Val Met Gly Ala Val
                325                 330                 335
Ile Met Glu Gly Phe Tyr Val Val Phe Asp Arg Ala Arg Lys Arg Ile
            340                 345                 350
Gly Phe Ala Val Ser Ala Cys His Val His Asp Glu Phe Arg Thr Ala
        355                 360                 365
Ala Val Glu Gly Pro Phe Val Thr Leu Asp Met Glu Asp Cys Gly Tyr
    370                 375                 380
Asn Ile Pro Gln Thr Asp Glu Ser Thr Leu Met Thr Ile Ala Tyr Val
385                 390                 395                 400
Met Ala Ala Ile Cys Ala Leu Phe Met Leu Pro Leu Cys Leu Met Val
                405                 410                 415
Cys Gln Trp Arg Cys Leu Arg Cys Leu Arg Gln Gln His Asp Asp Phe
            420                 425                 430
Ala Asp Asp Ile Ser Leu Leu Lys
        435                 440
```

<210> 140
<211> 810
<212> PRT
<213> Homo sapiens

<400> 140

```
Met Glu His Lys Glu Val Val Leu Leu Leu Leu Leu Phe Leu Lys Ser
1               5                   10                  15
Gly Gln Gly Glu Pro Leu Asp Asp Tyr Val Asn Thr Gln Gly Ala Ser
            20                  25                  30
Leu Phe Ser Val Thr Lys Lys Gln Leu Gly Ala Gly Ser Ile Glu Glu
        35                  40                  45
Cys Ala Ala Lys Cys Glu Glu Asp Glu Glu Phe Thr Cys Arg Ala Phe
    50                  55                  60
Gln Tyr His Ser Lys Glu Gln Gln Cys Val Ile Met Ala Glu Asn Arg
65                  70                  75                  80
Lys Ser Ser Ile Ile Ile Arg Met Arg Asp Val Val Leu Phe Glu Lys
                85                  90                  95
Lys Val Tyr Leu Ser Glu Cys Lys Thr Gly Asn Gly Lys Asn Tyr Arg
            100                 105                 110
Gly Thr Met Ser Lys Thr Lys Asn Gly Ile Thr Cys Gln Lys Trp Ser
        115                 120                 125
Ser Thr Ser Pro His Arg Pro Arg Phe Ser Pro Ala Thr His Pro Ser
    130                 135                 140
Glu Gly Leu Glu Glu Asn Tyr Cys Arg Asn Pro Asp Asn Asp Pro Gln
145                 150                 155                 160
Gly Pro Trp Cys Tyr Thr Thr Asp Pro Glu Lys Arg Tyr Asp Tyr Cys
                165                 170                 175
Asp Ile Leu Glu Cys Glu Glu Glu Cys Met His Cys Ser Gly Glu Asn
            180                 185                 190
Tyr Asp Gly Lys Ile Ser Lys Thr Met Ser Gly Leu Glu Cys Gln Ala
        195                 200                 205
Trp Asp Ser Gln Ser Pro His Ala His Gly Tyr Ile Pro Ser Lys Phe
    210                 215                 220
Pro Asn Lys Asn Leu Lys Lys Asn Tyr Cys Arg Asn Pro Asp Arg Glu
225                 230                 235                 240
Leu Arg Pro Trp Cys Phe Thr Thr Asp Pro Asn Lys Arg Trp Glu Leu
                245                 250                 255
Cys Asp Ile Pro Arg Cys Thr Thr Pro Pro Pro Ser Ser Gly Pro Thr
            260                 265                 270
```

Tyr Gln Cys Leu Lys Gly Thr Gly Glu Asn Tyr Arg Gly Asn Val Ala
275 280 285
Val Thr Val Ser Gly His Thr Cys Gln His Trp Ser Ala Gln Thr Pro
290 295 300
His Thr His Asn Arg Thr Pro Glu Asn Phe Pro Cys Lys Asn Leu Asp
305 310 315 320
Glu Asn Tyr Cys Arg Asn Pro Asp Gly Lys Arg Ala Pro Trp Cys His
325 330 335
Thr Thr Asn Ser Gln Val Arg Trp Glu Tyr Cys Lys Ile Pro Ser Cys
340 345 350
Asp Ser Ser Pro Val Ser Thr Glu Gln Leu Ala Pro Thr Ala Pro Pro
355 360 365
Glu Leu Thr Pro Val Val Gln Asp Cys Tyr His Gly Asp Gly Gln Ser
370 375 380
Tyr Arg Gly Thr Ser Ser Thr Thr Thr Thr Gly Lys Lys Cys Gln Ser
385 390 395 400
Trp Ser Ser Met Thr Pro His Arg His Gln Lys Thr Pro Glu Asn Tyr
405 410 415
Pro Asn Ala Gly Leu Thr Met Asn Tyr Cys Arg Asn Pro Asp Ala Asp
420 425 430
Lys Gly Pro Trp Cys Phe Thr Thr Asp Pro Ser Val Arg Trp Glu Tyr
435 440 445
Cys Asn Leu Lys Lys Cys Ser Gly Thr Glu Ala Ser Val Val Ala Pro
450 455 460
Pro Pro Val Val Leu Leu Pro Asp Val Glu Thr Pro Ser Glu Glu Asp
465 470 475 480
Cys Met Phe Gly Asn Gly Lys Gly Tyr Arg Gly Lys Arg Ala Thr Thr
485 490 495
Val Thr Gly Thr Pro Cys Gln Asp Trp Ala Ala Gln Glu Pro His Arg
500 505 510
His Ser Ile Phe Thr Pro Glu Thr Asn Pro Arg Ala Gly Leu Glu Lys
515 520 525
Asn Tyr Cys Arg Asn Pro Asp Gly Asp Val Gly Gly Pro Trp Cys Tyr
530 535 540
Thr Thr Asn Pro Arg Lys Leu Tyr Asp Tyr Cys Asp Val Pro Gln Cys
545 550 555 560
Ala Ala Pro Ser Phe Asp Cys Gly Lys Pro Gln Val Glu Pro Lys Lys
565 570 575
Cys Pro Gly Arg Val Val Gly Gly Cys Val Ala His Pro His Ser Trp
580 585 590
Pro Trp Gln Val Ser Leu Arg Thr Arg Phe Gly Met His Phe Cys Gly
595 600 605
Gly Thr Leu Ile Ser Pro Glu Trp Val Leu Thr Ala Ala His Cys Leu
610 615 620
Glu Lys Ser Pro Arg Pro Ser Ser Tyr Lys Val Ile Leu Gly Ala His
625 630 635 640
Gln Glu Val Asn Leu Glu Pro His Val Gln Glu Ile Glu Val Ser Arg
645 650 655
Leu Phe Leu Glu Pro Thr Arg Lys Asp Ile Ala Leu Leu Lys Leu Ser
660 665 670
Ser Pro Ala Val Ile Thr Asp Lys Val Ile Pro Ala Cys Leu Pro Ser
675 680 685
Pro Asn Tyr Val Val Ala Asp Arg Thr Glu Cys Phe Ile Thr Gly Trp
690 695 700
Gly Glu Thr Gln Gly Thr Phe Gly Ala Gly Leu Leu Lys Glu Ala Gln
705 710 715 720
Leu Pro Val Ile Glu Asn Lys Val Cys Asn Arg Tyr Glu Phe Leu Asn
725 730 735
Gly Arg Val Gln Ser Thr Glu Leu Cys Ala Gly His Leu Ala Gly Gly
740 745 750
Thr Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Cys Phe Glu
755 760 765
Lys Asp Lys Tyr Ile Leu Gln Gly Val Thr Ser Trp Gly Leu Gly Cys

```
    770                 775                 780
Ala Arg Pro Asn Lys Pro Gly Val Tyr Val Arg Val Ser Arg Phe Val
785                 790                 795                 800
Thr Trp Ile Glu Gly Val Met Arg Asn Asn
                805                 810
```

<210> 141
<211> 762
<212> PRT
<213> Homo sapiens

<400> 141

```
Gly Pro Asn Ile Cys Thr Thr Arg Gly Val Ser Ser Cys Gln Gln Cys
1               5                  10                  15
Leu Ala Val Ser Pro Met Cys Ala Trp Cys Ser Asp Glu Ala Leu Pro
            20                  25                  30
Leu Gly Ser Pro Arg Cys Asp Leu Lys Glu Asn Leu Leu Lys Asp Asn
        35                  40                  45
Cys Ala Pro Glu Ser Ile Glu Phe Pro Val Ser Glu Ala Arg Val Leu
    50                  55                  60
Glu Asp Arg Pro Leu Ser Asp Lys Gly Ser Gly Asp Ser Ser Gln Val
65                  70                  75                  80
Thr Gln Val Ser Pro Gln Arg Ile Ala Leu Arg Leu Arg Pro Asp Asp
                85                  90                  95
Ser Lys Asn Phe Ser Ile Gln Val Arg Gln Val Glu Asp Tyr Pro Val
            100                 105                 110
Asp Ile Tyr Tyr Leu Met Asp Leu Ser Tyr Ser Met Lys Asp Asp Leu
        115                 120                 125
Trp Ser Ile Gln Asn Leu Gly Thr Lys Leu Ala Thr Gln Met Arg Lys
    130                 135                 140
Leu Thr Ser Asn Leu Arg Ile Gly Phe Gly Ala Phe Val Asp Lys Pro
145                 150                 155                 160
Val Ser Pro Tyr Met Tyr Ile Ser Pro Pro Glu Ala Leu Glu Asn Pro
                165                 170                 175
Cys Tyr Asp Met Lys Thr Thr Cys Leu Pro Met Phe Gly Tyr Lys His
            180                 185                 190
Val Leu Thr Leu Thr Asp Gln Val Thr Arg Phe Asn Glu Glu Val Lys
        195                 200                 205
Lys Gln Ser Val Ser Arg Asn Arg Asp Ala Pro Glu Gly Gly Phe Asp
    210                 215                 220
Ala Ile Met Gln Ala Thr Val Cys Asp Glu Lys Ile Gly Trp Arg Asn
225                 230                 235                 240
Asp Ala Ser His Leu Leu Val Phe Thr Thr Asp Ala Lys Thr His Ile
                245                 250                 255
Ala Leu Asp Gly Arg Leu Ala Gly Ile Val Gln Pro Asn Asp Gly Gln
            260                 265                 270
Cys His Val Gly Ser Asp Asn His Tyr Ser Ala Ser Thr Thr Met Asp
        275                 280                 285
Tyr Pro Ser Leu Gly Leu Met Thr Glu Lys Leu Ser Gln Lys Asn Ile
    290                 295                 300
Asn Leu Ile Phe Ala Val Thr Glu Asn Val Val Asn Leu Tyr Gln Asn
305                 310                 315                 320
Tyr Ser Glu Leu Ile Pro Gly Thr Thr Val Gly Val Leu Ser Met Asp
                325                 330                 335
Ser Ser Asn Val Leu Gln Leu Ile Val Asp Ala Tyr Gly Lys Ile Arg
            340                 345                 350
Ser Lys Val Glu Leu Glu Val Arg Asp Leu Pro Glu Glu Leu Ser Leu
        355                 360                 365
Ser Phe Asn Ala Thr Cys Leu Asn Asn Glu Val Ile Pro Gly Leu Lys
    370                 375                 380
Ser Cys Met Gly Leu Lys Ile Gly Asp Thr Val Ser Phe Ser Ile Glu
385                 390                 395                 400
```

```
Ala Lys Val Arg Gly Cys Pro Gln Glu Lys Glu Lys Ser Phe Thr Ile
                405                 410                 415
Lys Pro Val Gly Phe Lys Asp Ser Leu Ile Val Gln Val Thr Phe Asp
            420                 425                 430
Cys Asp Cys Ala Cys Gln Ala Gln Ala Glu Pro Asn Ser His Arg Cys
        435                 440                 445
Asn Asn Gly Asn Gly Thr Phe Glu Cys Gly Val Cys Arg Cys Gly Pro
    450                 455                 460
Gly Trp Leu Gly Ser Gln Cys Glu Cys Ser Glu Glu Asp Tyr Arg Pro
465                 470                 475                 480
Ser Gln Gln Asp Glu Cys Ser Pro Arg Glu Gly Gln Pro Val Cys Ser
                485                 490                 495
Gln Arg Gly Glu Cys Leu Cys Gly Gln Cys Val Cys His Ser Ser Asp
            500                 505                 510
Phe Gly Lys Ile Thr Gly Lys Tyr Cys Glu Cys Asp Asp Phe Ser Cys
        515                 520                 525
Val Arg Tyr Lys Gly Glu Met Cys Ser Gly His Gly Gln Cys Ser Cys
    530                 535                 540
Gly Asp Cys Leu Cys Asp Ser Asp Trp Thr Gly Tyr Tyr Cys Asn Cys
545                 550                 555                 560
Thr Thr Arg Thr Asp Thr Cys Met Ser Ser Asn Gly Leu Leu Cys Ser
                565                 570                 575
Gly Arg Gly Lys Cys Glu Cys Gly Ser Cys Val Cys Ile Gln Pro Gly
            580                 585                 590
Ser Tyr Gly Asp Thr Cys Glu Lys Cys Pro Thr Cys Pro Asp Ala Cys
        595                 600                 605
Thr Phe Lys Lys Glu Cys Val Glu Cys Lys Lys Phe Asp Arg Glu Pro
    610                 615                 620
Tyr Met Thr Glu Asn Thr Cys Asn Arg Tyr Cys Arg Asp Glu Ile Glu
625                 630                 635                 640
Ser Val Lys Glu Leu Lys Asp Thr Gly Lys Asp Ala Val Asn Cys Thr
                645                 650                 655
Tyr Lys Asn Glu Asp Asp Cys Val Val Arg Phe Gln Tyr Tyr Glu Asp
            660                 665                 670
Ser Ser Gly Lys Ser Ile Leu Tyr Val Val Glu Glu Pro Glu Cys Pro
        675                 680                 685
Lys Gly Pro Asp Ile Leu Val Val Leu Leu Ser Val Met Gly Ala Ile
    690                 695                 700
Leu Leu Ile Gly Leu Ala Ala Leu Leu Ile Trp Lys Leu Leu Ile Thr
705                 710                 715                 720
Ile His Asp Arg Lys Glu Phe Ala Lys Phe Glu Glu Glu Arg Ala Arg
                725                 730                 735
Ala Lys Trp Asp Thr Ala Asn Asn Pro Leu Tyr Lys Glu Ala Thr Ser
            740                 745                 750
Thr Phe Thr Asn Ile Thr Tyr Arg Gly Thr
        755                 760
```

<210> 142
<211> 505
<212> PRT
<213> Homo sapiens

<400> 142

```
Gln Thr Ser Val Ser Pro Ser Lys Val Ile Leu Pro Arg Gly Gly Ser
1               5                   10                  15
Val Leu Val Thr Cys Ser Thr Ser Cys Asp Gln Pro Lys Leu Leu Gly
            20                  25                  30
Ile Glu Thr Pro Leu Pro Lys Lys Glu Leu Leu Leu Pro Gly Asn Asn
        35                  40                  45
Arg Lys Val Tyr Glu Leu Ser Asn Val Gln Glu Asp Ser Gln Pro Met
    50                  55                  60
Cys Tyr Ser Asn Cys Pro Asp Gly Gln Ser Thr Ala Lys Thr Phe Leu
```

65 70 75 80
Thr Val Tyr Trp Thr Pro Glu Arg Val Glu Leu Ala Pro Leu Pro Ser
85 90 95
Trp Gln Pro Val Gly Lys Asn Leu Thr Leu Arg Cys Gln Val Glu Gly
100 105 110
Gly Ala Pro Arg Ala Asn Leu Thr Val Val Leu Leu Arg Gly Glu Lys
115 120 125
Glu Leu Lys Arg Glu Pro Ala Val Gly Glu Pro Ala Glu Val Thr Thr
130 135 140
Thr Val Leu Val Arg Arg Asp His His Gly Ala Asn Phe Ser Cys Arg
145 150 155 160
Thr Glu Leu Asp Leu Arg Pro Gln Gly Leu Glu Leu Phe Glu Asn Thr
165 170 175
Ser Ala Pro Tyr Gln Leu Gln Thr Phe Val Leu Pro Ala Thr Pro Pro
180 185 190
Gln Leu Val Ser Pro Arg Val Leu Glu Val Asp Thr Gln Gly Thr Val
195 200 205
Val Cys Ser Leu Asp Gly Leu Phe Pro Val Ser Glu Ala Gln Val His
210 215 220
Leu Ala Leu Gly Asp Gln Arg Leu Asn Pro Thr Val Thr Tyr Gly Asn
225 230 235 240
Asp Ser Phe Ser Ala Lys Ala Ser Val Ser Val Thr Ala Glu Asp Glu
245 250 255
Gly Thr Gln Arg Leu Thr Cys Ala Val Ile Leu Gly Asn Gln Ser Gln
260 265 270
Glu Thr Leu Gln Thr Val Thr Ile Tyr Ser Phe Pro Ala Pro Asn Val
275 280 285
Ile Leu Thr Lys Pro Glu Val Ser Glu Gly Thr Glu Val Thr Val Lys
290 295 300
Cys Glu Ala His Pro Arg Ala Lys Val Thr Leu Asn Gly Val Pro Ala
305 310 315 320
Gln Pro Leu Gly Pro Arg Ala Gln Leu Leu Leu Lys Ala Thr Pro Glu
325 330 335
Asp Asn Gly Arg Ser Phe Ser Cys Ser Ala Thr Leu Glu Val Ala Gly
340 345 350
Gln Leu Ile His Lys Asn Gln Thr Arg Glu Leu Arg Val Leu Tyr Gly
355 360 365
Pro Arg Leu Asp Glu Arg Asp Cys Pro Gly Asn Trp Thr Trp Pro Glu
370 375 380
Asn Ser Gln Gln Thr Pro Met Cys Gln Ala Trp Gly Asn Pro Leu Pro
385 390 395 400
Glu Leu Lys Cys Leu Lys Asp Gly Thr Phe Pro Leu Pro Ile Gly Glu
405 410 415
Ser Val Thr Val Thr Arg Asp Leu Glu Gly Thr Tyr Leu Cys Arg Ala
420 425 430
Arg Ser Thr Gln Gly Glu Val Thr Arg Glu Val Thr Val Asn Val Leu
435 440 445
Ser Pro Arg Tyr Glu Ile Val Ile Ile Thr Val Val Ala Ala Ala Val
450 455 460
Ile Met Gly Thr Ala Gly Leu Ser Thr Tyr Leu Tyr Asn Arg Gln Arg
465 470 475 480
Lys Ile Lys Lys Tyr Arg Leu Gln Gln Ala Gln Lys Gly Thr Pro Met
485 490 495
Lys Pro Asn Thr Gln Ala Thr Pro Pro
500 505

<210> 143
<211> 261
<212> PRT
<213> Homo sapiens

<400> 143

```
Met Ile Glu Thr Tyr Asn Gln Thr Ser Pro Arg Ser Ala Ala Thr Gly
1               5                  10                  15
Leu Pro Ile Ser Met Lys Ile Phe Met Tyr Leu Leu Thr Val Phe Leu
            20                  25                  30
Ile Thr Gln Met Ile Gly Ser Ala Leu Phe Ala Val Tyr Leu His Arg
        35                  40                  45
Arg Leu Asp Lys Ile Glu Asp Glu Arg Asn Leu His Glu Asp Phe Val
    50                  55                  60
Phe Met Lys Thr Ile Gln Arg Cys Asn Thr Gly Glu Arg Ser Leu Ser
65                  70                  75                  80
Leu Leu Asn Cys Glu Glu Ile Lys Ser Gln Phe Glu Gly Phe Val Lys
                85                  90                  95
Asp Ile Met Leu Asn Lys Glu Glu Thr Lys Lys Glu Asn Ser Phe Glu
            100                 105                 110
Met Gln Lys Gly Asp Gln Asn Pro Gln Ile Ala Ala His Val Ile Ser
        115                 120                 125
Glu Ala Ser Ser Lys Thr Thr Ser Val Leu Gln Trp Ala Glu Lys Gly
    130                 135                 140
Tyr Tyr Thr Met Ser Asn Asn Leu Val Thr Leu Glu Asn Gly Lys Gln
145                 150                 155                 160
Leu Thr Val Lys Arg Gln Gly Leu Tyr Tyr Ile Tyr Ala Gln Val Thr
                165                 170                 175
Phe Cys Ser Asn Arg Glu Ala Ser Ser Gln Ala Pro Phe Ile Ala Ser
            180                 185                 190
Leu Cys Leu Lys Ser Pro Gly Arg Phe Glu Arg Ile Leu Leu Arg Ala
        195                 200                 205
Ala Asn Thr His Ser Ser Ala Lys Pro Cys Gly Gln Gln Ser Ile His
    210                 215                 220
Leu Gly Gly Val Phe Glu Leu Gln Pro Gly Ala Ser Val Phe Val Asn
225                 230                 235                 240
Val Thr Asp Pro Ser Gln Val Ser His Gly Thr Gly Phe Thr Ser Phe
                245                 250                 255
Gly Leu Leu Lys Leu
            260
```

<210> 144
<211> 187
<212> PRT
<213> Homo sapiens

<400> 144

```
Ala Met His Val Ala Gln Pro Ala Val Val Leu Ala Ser Ser Arg Gly
1               5                  10                  15
Ile Ala Ser Phe Val Cys Glu Tyr Ala Ser Pro Gly Lys Ala Thr Glu
            20                  25                  30
Val Arg Val Thr Val Leu Arg Gln Ala Asp Ser Gln Val Thr Glu Val
        35                  40                  45
Cys Ala Ala Thr Tyr Met Met Gly Asn Glu Leu Thr Phe Leu Asp Asp
    50                  55                  60
Ser Ile Cys Thr Gly Thr Ser Ser Gly Asn Gln Val Asn Leu Thr Ile
65                  70                  75                  80
Gln Gly Leu Arg Ala Met Asp Thr Gly Leu Tyr Ile Cys Lys Val Glu
                85                  90                  95
Leu Met Tyr Pro Pro Pro Tyr Tyr Leu Gly Ile Gly Asn Gly Thr Gln
            100                 105                 110
Ile Tyr Val Ile Asp Pro Glu Pro Cys Pro Asp Ser Asp Phe Leu Leu
        115                 120                 125
Trp Ile Leu Ala Ala Val Ser Ser Gly Leu Phe Phe Tyr Ser Phe Leu
    130                 135                 140
Leu Thr Ala Val Ser Leu Ser Lys Met Leu Lys Lys Arg Ser Pro Leu
145                 150                 155                 160
Thr Thr Gly Val Tyr Val Lys Met Pro Pro Thr Glu Pro Glu Cys Glu
```

```
                 165                 170                 175
Lys Gln Phe Gln Pro Tyr Phe Ile Pro Ile Asn
            180                 185
```

<210> 145
<211> 544
<212> PRT
<213> Homo sapiens

<400> 145

```
Ile Pro Pro His Val Gln Lys Ser Val Asn Asn Asp Met Ile Val Thr
1               5                  10                  15
Asp Asn Asn Gly Ala Val Lys Phe Pro Gln Leu Cys Lys Phe Cys Asp
            20                  25                  30
Val Arg Phe Ser Thr Cys Asp Asn Gln Lys Ser Cys Met Ser Asn Cys
        35                  40                  45
Ser Ile Thr Ser Ile Cys Glu Lys Pro Gln Glu Val Cys Val Ala Val
    50                  55                  60
Trp Arg Lys Asn Asp Glu Asn Ile Thr Leu Glu Thr Val Cys His Asp
65                  70                  75                  80
Pro Lys Leu Pro Tyr His Asp Phe Ile Leu Glu Asp Ala Ala Ser Pro
                85                  90                  95
Lys Cys Ile Met Lys Glu Lys Lys Lys Pro Gly Glu Thr Phe Phe Met
            100                 105                 110
Cys Ser Cys Ser Ser Asp Glu Cys Asn Asp Asn Ile Ile Phe Ser Glu
        115                 120                 125
Glu Tyr Asn Thr Ser Asn Pro Asp Leu Leu Leu Val Ile Phe Gln Val
    130                 135                 140
Thr Gly Ile Ser Leu Leu Pro Pro Leu Gly Val Ala Ile Ser Val Ile
145                 150                 155                 160
Ile Ile Phe Tyr Cys Tyr Arg Val Asn Arg Gln Gln Lys Leu Ser Ser
                165                 170                 175
Thr Trp Glu Thr Gly Lys Thr Arg Lys Leu Met Glu Phe Ser Glu His
            180                 185                 190
Cys Ala Ile Ile Leu Glu Asp Asp Arg Ser Asp Ile Ser Ser Thr Cys
        195                 200                 205
Ala Asn Asn Ile Asn His Asn Thr Glu Leu Leu Pro Ile Glu Leu Asp
    210                 215                 220
Thr Leu Val Gly Lys Gly Arg Phe Ala Glu Val Tyr Lys Ala Lys Leu
225                 230                 235                 240
Lys Gln Asn Thr Ser Glu Gln Phe Glu Thr Val Ala Val Lys Ile Phe
                245                 250                 255
Pro Tyr Glu Glu Tyr Ala Ser Trp Lys Thr Glu Lys Asp Ile Phe Ser
            260                 265                 270
Asp Ile Asn Leu Lys His Glu Asn Ile Leu Gln Phe Leu Thr Ala Glu
        275                 280                 285
Glu Arg Lys Thr Glu Leu Gly Lys Gln Tyr Trp Leu Ile Thr Ala Phe
    290                 295                 300
His Ala Lys Gly Asn Leu Gln Glu Tyr Leu Thr Arg His Val Ile Ser
305                 310                 315                 320
Trp Glu Asp Leu Arg Lys Leu Gly Ser Ser Leu Ala Arg Gly Ile Ala
                325                 330                 335
His Leu His Ser Asp His Thr Pro Cys Gly Arg Pro Lys Met Pro Ile
            340                 345                 350
Val His Arg Asp Leu Lys Ser Ser Asn Ile Leu Val Lys Asn Asp Leu
        355                 360                 365
Thr Cys Cys Leu Cys Asp Phe Gly Leu Ser Leu Arg Leu Asp Pro Thr
    370                 375                 380
Leu Ser Val Asp Asp Leu Ala Asn Ser Gly Gln Val Gly Thr Ala Arg
385                 390                 395                 400
Tyr Met Ala Pro Glu Val Leu Glu Ser Arg Met Asn Leu Glu Asn Ala
                405                 410                 415
```

```
Glu Ser Phe Lys Gln Thr Asp Val Tyr Ser Met Ala Leu Val Leu Trp
            420                 425                 430
Glu Met Thr Ser Arg Cys Asn Ala Val Gly Glu Val Lys Asp Tyr Glu
        435                 440                 445
Pro Pro Phe Gly Ser Lys Val Arg Glu His Pro Cys Val Glu Ser Met
    450                 455                 460
Lys Asp Asn Val Leu Arg Asp Arg Gly Arg Pro Glu Ile Pro Ser Phe
465                 470                 475                 480
Trp Leu Asn His Gln Gly Ile Gln Met Val Cys Glu Thr Leu Thr Glu
                485                 490                 495
Cys Trp Asp His Asp Pro Glu Ala Arg Leu Thr Ala Gln Cys Val Ala
            500                 505                 510
Glu Arg Phe Ser Glu Leu Glu His Leu Asp Arg Leu Ser Gly Arg Ser
        515                 520                 525
Cys Ser Glu Glu Lys Ile Pro Glu Asp Gly Ser Leu Asn Thr Thr Lys
    530                 535                 540
```

<210> 146
<211> 358
<212> PRT
<213> Homo sapiens

<400> 146

```
Cys Glu Glu Pro Pro Thr Phe Glu Ala Met Glu Leu Ile Gly Lys Pro
1               5                   10                  15
Lys Pro Tyr Tyr Glu Ile Gly Glu Arg Val Asp Tyr Lys Cys Lys Lys
            20                  25                  30
Gly Tyr Phe Tyr Ile Pro Pro Leu Ala Thr His Thr Ile Cys Asp Arg
        35                  40                  45
Asn His Thr Trp Leu Pro Val Ser Asp Asp Ala Cys Tyr Arg Glu Thr
    50                  55                  60
Cys Pro Tyr Ile Arg Asp Pro Leu Asn Gly Gln Ala Val Pro Ala Asn
65                  70                  75                  80
Gly Thr Tyr Glu Phe Gly Tyr Gln Met His Phe Ile Cys Asn Glu Gly
                85                  90                  95
Tyr Tyr Leu Ile Gly Glu Glu Ile Leu Tyr Cys Glu Leu Lys Gly Ser
            100                 105                 110
Val Ala Ile Trp Ser Gly Lys Pro Pro Ile Cys Glu Lys Val Leu Cys
        115                 120                 125
Thr Pro Pro Pro Lys Ile Lys Asn Gly Lys His Thr Phe Ser Glu Val
    130                 135                 140
Glu Val Phe Glu Tyr Leu Asp Ala Val Thr Tyr Ser Cys Asp Pro Ala
145                 150                 155                 160
Pro Gly Pro Asp Pro Phe Ser Leu Ile Gly Glu Ser Thr Ile Tyr Cys
                165                 170                 175
Gly Asp Asn Ser Val Trp Ser Arg Ala Ala Pro Glu Cys Lys Val Val
            180                 185                 190
Lys Cys Arg Phe Pro Val Val Glu Asn Gly Lys Gln Ile Ser Gly Phe
        195                 200                 205
Gly Lys Lys Phe Tyr Tyr Lys Ala Thr Val Met Phe Glu Cys Asp Lys
    210                 215                 220
Gly Phe Tyr Leu Asp Gly Ser Asp Thr Ile Val Cys Asp Ser Asn Ser
225                 230                 235                 240
Thr Trp Asp Pro Pro Val Pro Lys Cys Leu Lys Val Leu Pro Pro Ser
                245                 250                 255
Ser Thr Lys Pro Pro Ala Leu Ser His Ser Val Ser Thr Ser Ser Thr
            260                 265                 270
Thr Lys Ser Pro Ala Ser Ser Ala Ser Gly Pro Arg Pro Thr Tyr Lys
        275                 280                 285
Pro Pro Val Ser Asn Tyr Pro Gly Tyr Pro Lys Pro Glu Glu Gly Ile
    290                 295                 300
Leu Asp Ser Leu Asp Val Trp Val Ile Ala Val Ile Val Ile Ala Ile
```

305 310 315 320
Val Val Gly Val Ala Val Ile Cys Val Val Pro Tyr Arg Tyr Leu Gln
325 330 335
Arg Arg Lys Lys Lys Gly Thr Tyr Leu Thr Asp Glu Thr His Arg Glu
340 345 350
Val Lys Phe Thr Ser Leu
355

<210> 147
<211> 1148
<212> PRT
<213> Homo sapiens

<400> 147
Leu Pro Glu Ala Lys Ile Phe Ser Gly Pro Ser Ser Glu Gln Phe Gly
1 5 10 15
Tyr Ala Val Gln Gln Phe Ile Asn Pro Lys Gly Asn Trp Leu Leu Val
20 25 30
Gly Ser Pro Trp Ser Gly Phe Pro Glu Asn Arg Met Gly Asp Val Tyr
35 40 45
Lys Cys Pro Val Asp Leu Ser Thr Ala Thr Cys Glu Lys Leu Asn Leu
50 55 60
Gln Thr Ser Thr Ser Ile Pro Asn Val Thr Glu Met Lys Thr Asn Met
65 70 75 80
Ser Leu Gly Leu Ile Leu Thr Arg Asn Met Gly Thr Gly Gly Phe Leu
85 90 95
Thr Cys Gly Pro Leu Trp Ala Gln Gln Cys Gly Asn Gln Tyr Tyr Thr
100 105 110
Thr Gly Val Cys Ser Asp Ile Ser Pro Asp Phe Gln Leu Ser Ala Ser
115 120 125
Phe Ser Pro Ala Thr Gln Pro Cys Pro Ser Leu Ile Asp Val Val Val
130 135 140
Val Cys Asp Glu Ser Asn Ser Ile Tyr Pro Trp Asp Ala Val Lys Asn
145 150 155 160
Phe Leu Glu Lys Phe Val Gln Gly Leu Asp Ile Gly Pro Thr Lys Thr
165 170 175
Gln Val Gly Leu Ile Gln Tyr Ala Asn Asn Pro Arg Val Val Phe Asn
180 185 190
Leu Asn Thr Tyr Lys Thr Lys Glu Glu Met Ile Val Ala Thr Ser Gln
195 200 205
Thr Ser Gln Tyr Gly Gly Asp Leu Thr Asn Thr Phe Gly Ala Ile Gln
210 215 220
Tyr Ala Arg Lys Tyr Ala Tyr Ser Ala Ala Ser Gly Gly Arg Arg Ser
225 230 235 240
Ala Thr Lys Val Met Val Val Val Thr Asp Gly Glu Ser His Asp Gly
245 250 255
Ser Met Leu Lys Ala Val Ile Asp Gln Cys Asn His Asp Asn Ile Leu
260 265 270
Arg Phe Gly Ile Ala Val Leu Gly Tyr Leu Asn Arg Asn Ala Leu Asp
275 280 285
Thr Lys Asn Leu Ile Lys Glu Ile Lys Ala Ile Ala Ser Ile Pro Thr
290 295 300
Glu Arg Tyr Phe Phe Asn Val Ser Asp Glu Ala Ala Leu Leu Glu Lys
305 310 315 320
Ala Gly Thr Leu Gly Glu Gln Ile Phe Ser Ile Glu Gly Thr Val Gln
325 330 335
Gly Gly Asp Asn Phe Gln Met Glu Met Ser Gln Val Gly Phe Ser Ala
340 345 350
Asp Tyr Ser Ser Gln Asn Asp Ile Leu Met Leu Gly Ala Val Gly Ala
355 360 365
Phe Gly Trp Ser Gly Thr Ile Val Gln Lys Thr Ser His Gly His Leu
370 375 380
Ile Phe Pro Lys Gln Ala Phe Asp Gln Ile Leu Gln Asp Arg Asn His

385 390 395 400
Ser Ser Tyr Leu Gly Tyr Ser Val Ala Ala Ile Ser Thr Gly Glu Ser
405 410 415
Thr His Phe Val Ala Gly Ala Pro Arg Ala Asn Tyr Thr Gly Gln Ile
420 425 430
Val Leu Tyr Ser Val Asn Glu Asn Gly Asn Ile Thr Val Ile Gln Ala
435 440 445
His Arg Gly Asp Gln Ile Gly Ser Tyr Phe Gly Ser Val Leu Cys Ser
450 455 460
Val Asp Val Asp Lys Asp Thr Ile Thr Asp Val Leu Leu Val Gly Ala
465 470 475 480
Pro Met Tyr Met Ser Asp Leu Lys Lys Glu Glu Gly Arg Val Tyr Leu
485 490 495
Phe Thr Ile Lys Lys Gly Ile Leu Gly Gln His Gln Phe Leu Glu Gly
500 505 510
Pro Glu Gly Ile Glu Asn Thr Arg Phe Gly Ser Ala Ile Ala Ala Leu
515 520 525
Ser Asp Ile Asn Met Asp Gly Phe Asn Asp Val Ile Val Gly Ser Pro
530 535 540
Leu Glu Asn Gln Asn Ser Gly Ala Val Tyr Ile Tyr Asn Gly His Gln
545 550 555 560
Gly Thr Ile Arg Thr Lys Tyr Ser Gln Lys Ile Leu Gly Ser Asp Gly
565 570 575
Ala Phe Arg Ser His Leu Gln Tyr Phe Gly Arg Ser Leu Asp Gly Tyr
580 585 590
Gly Asp Leu Asn Gly Asp Ser Ile Thr Asp Val Ser Ile Gly Ala Phe
595 600 605
Gly Gln Val Val Gln Leu Trp Ser Gln Ser Ile Ala Asp Val Ala Ile
610 615 620
Glu Ala Ser Phe Thr Pro Glu Lys Ile Thr Leu Val Asn Lys Asn Ala
625 630 635 640
Gln Ile Ile Leu Lys Leu Cys Phe Ser Ala Lys Phe Arg Pro Thr Lys
645 650 655
Gln Asn Asn Gln Val Ala Ile Val Tyr Asn Ile Thr Leu Asp Ala Asp
660 665 670
Gly Phe Ser Ser Arg Val Thr Ser Arg Gly Leu Phe Lys Glu Asn Asn
675 680 685
Glu Arg Cys Leu Gln Lys Asn Met Val Val Asn Gln Ala Gln Ser Cys
690 695 700
Pro Glu His Ile Ile Tyr Ile Gln Glu Pro Ser Asp Val Val Asn Ser
705 710 715 720
Leu Asp Leu Arg Val Asp Ile Ser Leu Glu Asn Pro Gly Thr Ser Pro
725 730 735
Ala Leu Glu Ala Tyr Ser Glu Thr Ala Lys Val Phe Ser Ile Pro Phe
740 745 750
His Lys Asp Cys Gly Glu Asp Gly Leu Cys Ile Ser Asp Leu Val Leu
755 760 765
Asp Val Arg Gln Ile Pro Ala Ala Gln Glu Gln Pro Phe Ile Val Ser
770 775 780
Asn Gln Asn Lys Arg Leu Thr Phe Ser Val Thr Leu Lys Asn Lys Arg
785 790 795 800
Glu Ser Ala Tyr Asn Thr Gly Ile Val Val Asp Phe Ser Glu Asn Leu
805 810 815
Phe Phe Ala Ser Phe Ser Leu Pro Val Asp Gly Thr Glu Val Thr Cys
820 825 830
Gln Val Ala Ala Ser Gln Lys Ser Val Ala Cys Asp Val Gly Tyr Pro
835 840 845
Ala Leu Lys Arg Glu Gln Gln Val Thr Phe Thr Ile Asn Phe Asp Phe
850 855 860
Asn Leu Gln Asn Leu Gln Asn Gln Ala Ser Leu Ser Phe Gln Ala Leu
865 870 875 880
Ser Glu Ser Gln Glu Glu Asn Lys Ala Asp Asn Leu Val Asn Leu Lys
885 890 895

```
Ile Pro Leu Leu Tyr Asp Ala Glu Ile His Leu Thr Arg Ser Thr Asn
            900                 905                 910
Ile Asn Phe Tyr Glu Ile Ser Ser Asp Gly Asn Val Pro Ser Ile Val
        915                 920                 925
His Ser Phe Glu Asp Val Gly Pro Lys Phe Ile Phe Ser Leu Lys Val
    930                 935                 940
Thr Thr Gly Ser Val Pro Val Ser Met Ala Thr Val Ile Ile His Ile
945                 950                 955                 960
Pro Gln Tyr Thr Lys Glu Lys Asn Pro Leu Met Tyr Leu Thr Gly Val
                965                 970                 975
Gln Thr Asp Lys Ala Gly Asp Ile Ser Cys Asn Ala Asp Ile Asn Pro
            980                 985                 990
Leu Lys Ile Gly Gln Thr Ser Ser  Ser Val Ser Phe Lys  Ser Glu Asn
        995                 1000                1005
Phe Arg  His Thr Lys Glu Leu  Asn Cys Arg Thr Ala  Ser Cys Ser
    1010                1015                1020
Asn Val  Thr Cys Trp Leu Lys  Asp Val His Met Lys  Gly Glu Tyr
    1025                1030                1035
Phe Val  Asn Val Thr Thr Arg  Ile Trp Asn Gly Thr  Phe Ala Ser
    1040                1045                1050
Ser Thr  Phe Gln Thr Val Gln  Leu Thr Ala Ala Ala  Glu Ile Asn
    1055                1060                1065
Thr Tyr  Asn Pro Glu Ile Tyr  Val Ile Glu Asp Asn  Thr Val Thr
    1070                1075                1080
Ile Pro  Leu Met Ile Met Lys  Pro Asp Glu Lys Ala  Glu Val Pro
    1085                1090                1095
Thr Gly  Val Ile Ile Gly Ser  Ile Ile Ala Gly Ile  Leu Leu Leu
    1100                1105                1110
Leu Ala  Leu Val Ala Ile Leu  Trp Lys Leu Gly Phe  Phe Lys Arg
    1115                1120                1125
Lys Tyr  Glu Lys Met Thr Lys  Asn Pro Asp Glu Ile  Asp Glu Thr
    1130                1135                1140
Thr Glu  Leu Ser Ser
    1145
```

<210> 148
<211> 133
<212> PRT
<213> Homo sapiens

<400> 148

```
Ala Pro Met Thr Gln Thr Thr Pro Leu Lys Thr Ser Trp Val Asn Cys
1               5                  10                  15
Ser Asn Met Ile Asp Glu Ile Ile Thr His Leu Lys Gln Pro Pro Leu
            20                  25                  30
Pro Leu Leu Asp Phe Asn Asn Leu Asn Gly Glu Asp Gln Asp Ile Leu
        35                  40                  45
Met Glu Asn Asn Leu Arg Arg Pro Asn Leu Glu Ala Phe Asn Arg Ala
    50                  55                  60
Val Lys Ser Leu Gln Asn Ala Ser Ala Ile Glu Ser Ile Leu Lys Asn
65                  70                  75                  80
Leu Leu Pro Cys Leu Pro Leu Ala Thr Ala Ala Pro Thr Arg His Pro
                85                  90                  95
Ile His Ile Lys Asp Gly Asp Trp Asn Glu Phe Arg Arg Lys Leu Thr
            100                 105                 110
Phe Tyr Leu Lys Thr Leu Glu Asn Ala Gln Ala Gln Gln Thr Thr Leu
        115                 120                 125
Ser Leu Ala Ile Phe
    130
```

<210> 149
<211> 622
<212> PRT
<213> Homo sapiens

<400> 149

```
Met Ala His Val Arg Gly Leu Gln Leu Pro Gly Cys Leu Ala Leu Ala
1               5                  10                  15
Ala Leu Cys Ser Leu Val His Ser Gln His Val Phe Leu Ala Pro Gln
            20                  25                  30
Gln Ala Arg Ser Leu Leu Gln Arg Val Arg Arg Ala Asn Thr Phe Leu
        35                  40                  45
Glu Glu Val Arg Lys Gly Asn Leu Glu Arg Glu Cys Val Glu Glu Thr
    50                  55                  60
Cys Ser Tyr Glu Glu Ala Phe Glu Ala Leu Glu Ser Ser Thr Ala Thr
65                  70                  75                  80
Asp Val Phe Trp Ala Lys Tyr Thr Ala Cys Glu Thr Ala Arg Thr Pro
                85                  90                  95
Arg Asp Lys Leu Ala Ala Cys Leu Glu Gly Asn Cys Ala Glu Gly Leu
            100                 105                 110
Gly Thr Asn Tyr Arg Gly His Val Asn Ile Thr Arg Ser Gly Ile Glu
        115                 120                 125
Cys Gln Leu Trp Arg Ser Arg Tyr Pro His Lys Pro Glu Ile Asn Ser
    130                 135                 140
Thr Thr His Pro Gly Ala Asp Leu Gln Glu Asn Phe Cys Arg Asn Pro
145                 150                 155                 160
Asp Ser Ser Thr Thr Gly Pro Trp Cys Tyr Thr Thr Asp Pro Thr Val
                165                 170                 175
Arg Arg Gln Glu Cys Ser Ile Pro Val Cys Gly Gln Asp Gln Val Thr
            180                 185                 190
Val Ala Met Thr Pro Arg Ser Glu Gly Ser Ser Val Asn Leu Ser Pro
        195                 200                 205
Pro Leu Glu Gln Cys Val Pro Asp Arg Gly Gln Gln Tyr Gln Gly Arg
    210                 215                 220
Leu Ala Val Thr Thr His Gly Leu Pro Cys Leu Ala Trp Ala Ser Ala
225                 230                 235                 240
Gln Ala Lys Ala Leu Ser Lys His Gln Asp Phe Asn Ser Ala Val Gln
                245                 250                 255
Leu Val Glu Asn Phe Cys Arg Asn Pro Asp Gly Asp Glu Glu Gly Val
            260                 265                 270
Trp Cys Tyr Val Ala Gly Lys Pro Gly Asp Phe Gly Tyr Cys Asp Leu
        275                 280                 285
Asn Tyr Cys Glu Glu Ala Val Glu Glu Glu Thr Gly Asp Gly Leu Asp
    290                 295                 300
Glu Asp Ser Asp Arg Ala Ile Glu Gly Arg Thr Ala Thr Ser Glu Tyr
305                 310                 315                 320
Gln Thr Phe Phe Asn Pro Arg Thr Phe Gly Ser Gly Glu Ala Asp Cys
                325                 330                 335
Gly Leu Arg Pro Leu Phe Glu Lys Lys Ser Leu Glu Asp Lys Thr Glu
            340                 345                 350
Arg Glu Leu Leu Glu Ser Tyr Ile Asp Gly Arg Ile Val Glu Gly Ser
        355                 360                 365
Asp Ala Glu Ile Gly Met Ser Pro Trp Gln Val Met Leu Phe Arg Lys
    370                 375                 380
Ser Pro Gln Glu Leu Leu Cys Gly Ala Ser Leu Ile Ser Asp Arg Trp
385                 390                 395                 400
Val Leu Thr Ala Ala His Cys Leu Leu Tyr Pro Pro Trp Asp Lys Asn
                405                 410                 415
Phe Thr Glu Asn Asp Leu Leu Val Arg Ile Gly Lys His Ser Arg Thr
            420                 425                 430
Arg Tyr Glu Arg Asn Ile Glu Lys Ile Ser Met Leu Glu Lys Ile Tyr
        435                 440                 445
Ile His Pro Arg Tyr Asn Trp Arg Glu Asn Leu Asp Arg Asp Ile Ala
```

```
    450                 455                 460
Leu Met Lys Leu Lys Lys Pro Val Ala Phe Ser Asp Tyr Ile His Pro
465                 470                 475                 480
Val Cys Leu Pro Asp Arg Glu Thr Ala Ala Ser Leu Leu Gln Ala Gly
                485                 490                 495
Tyr Lys Gly Arg Val Thr Gly Trp Gly Asn Leu Lys Glu Thr Trp Thr
            500                 505                 510
Ala Asn Val Gly Lys Gly Gln Pro Ser Val Leu Gln Val Val Asn Leu
        515                 520                 525
Pro Ile Val Glu Arg Pro Val Cys Lys Asp Ser Thr Arg Ile Arg Ile
    530                 535                 540
Thr Asp Asn Met Phe Cys Ala Gly Tyr Lys Pro Asp Glu Gly Lys Arg
545                 550                 555                 560
Gly Asp Ala Cys Glu Gly Asp Ser Gly Gly Pro Phe Val Met Lys Ser
                565                 570                 575
Pro Phe Asn Asn Arg Trp Tyr Gln Met Gly Ile Val Ser Trp Gly Glu
            580                 585                 590
Gly Cys Asp Arg Asp Gly Lys Tyr Gly Phe Tyr Thr His Val Phe Arg
        595                 600                 605
Leu Lys Lys Trp Ile Gln Lys Val Ile Asp Gln Phe Gly Glu
    610                 615                 620
```

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/051173

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N9/56 C12N9/50 C12N9/64 C12N9/00 A61K38/48

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, EMBASE, WPI Data, PAJ, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	WO 00/78332 A (BJARNASON JON BRAGI) 28 December 2000 (2000-12-28) page 8, line 6 - page 11, line 5 -----	1-67, 77-82
Y	AGGARWAL B B ET AL: "HUMAN TUMOR NECROSIS FACTOR PRODUCTION, PURIFICATION, AND CHARACTERIZATION" JOURNAL OF BIOLOGICAL CHEMISTRY, THE AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, INC.,, US, vol. 260, no. 4, 25 February 1985 (1985-02-25), pages 2345-2354, XP000654946 ISSN: 0021-9258 page 2345, left-hand column, line 1 - page 2345, right-hand column, line 4 figure 7 ----- -/--	1-67, 77-82

[X] Further documents are listed in the continuation of box C

[X] Patent family members are listed in annex

° Special categories of cited documents

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 November 2004	08/12/2004

Name and mailing address of the ISA	Authorized officer
European Patent Office, P B 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel (+31-70) 340-2040, Tx 31 651 epo nl Fax (+31-70) 340-3016	Seroz, T

Form PCT/ISA/210 (second sheet) (January 2004)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passage	Relevant to claim No
Y	VAN KESSEL K P ET AL: "Inactivation of recombinant human tumor necrosis factor-alpha by proteolytic enzymes released from stimulated human neutrophils." JOURNAL OF IMMUNOLOGY (BALTIMORE, MD. : 1950) 1 DEC 1991, vol. 147, no. 11, 1 December 1991 (1991-12-01), pages 3862-3868, XP002304822 ISSN: 0022-1767 page 3867, left-hand column, paragraph 2; figure 2	1-67, 77-82
Y	XU Y ET AL: "Mutational analysis of the primary substrate specificity pocket of complement factor B. Asp(226) is a major structural determinant for p(1)-Arg binding." THE JOURNAL OF BIOLOGICAL CHEMISTRY. 7 JAN 2000, vol. 275, no. 1, 7 January 2000 (2000-01-07), pages 378-385, XP002304823 ISSN: 0021-9258 page 381, left-hand column, last paragraph - page 382, left-hand column, paragraph 1 page 385, left-hand column	1,2, 68-71, 79-82
Y	FAUTREL B ET AL: "INTERET DES MOLECULES ANTI-TNF-ALPHA DANS LE MALADIES INFLAMMATOIRES ET INFECTIEUSES" REVUE DE MEDECINE INTERNE, CMR, ASNIERES, FR, vol. 21, no. 10, 2000, pages 872-888, XP000965586 ISSN: 0248-8663 page 873, left-hand column, paragraph 1	1-67, 77-82
P,Y	RUGGLES SANDRA WAUGH ET AL: "Characterization of structural determinants of granzyme B reveals potent mediators of extended substrate specificity." THE JOURNAL OF BIOLOGICAL CHEMISTRY. 16 JUL 2004, vol. 279, no. 29, 16 July 2004 (2004-07-16), pages 30751-30759, XP002304824 ISSN: 0021-9258 page 30751, right-hand column, paragraph 1 figure 5	1,2, 68-71, 79-82

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No
A	ALTAMIRANO M M ET AL: "Directed evolution of new catalytic activity using the alpha/beta-barrel scaffold" NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, vol. 403, no. 6770, 10 February 2000 (2000-02-10), pages 617-622, XP002173865 ISSN: 0028-0836 page 618, right-hand column, last paragraph - page 619, left-hand column, paragraph 2; figure 3 -----	1-82

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2004/051173

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.
because they relate to subject matter not required to be searched by this Authority, namely

see FURTHER INFORMATION sheet PCT/ISA/210

2. [X] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest

[ ] No protest accompanied the payment of additional search fees

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box II.1

Although claim 79 is directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound.

Although claims 81 and 82 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound.

Continuation of Box II.2

Present claims 1-5, 7-71,74-82 relate to a product defined by reference to a desirable characteristic or property, namely, the capability to hydrolyse defined substrates at defined positions.
An attempt is made to define the product by reference to a result to be achieved. This lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the engineered trypsin, subtilisin E, human pesin and human caspase 7.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.5), should the problems which led to the Article 17(2) declaration be overcome.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/051173

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0078332 A	28-12-2000	AT 278417 T	15-10-2004
		AU 4947800 A	09-01-2001
		CA 2377357 A1	28-12-2000
		CN 1356907 T	03-07-2002
		EP 1202743 A2	08-05-2002
		WO 0078332 A2	28-12-2000
		JP 2003502071 T	21-01-2003
		NO 20016159 A	17-12-2001
		NZ 516632 A	30-04-2004
		PL 352318 A1	11-08-2003
		US 2002141987 A1	03-10-2002